The Dissertation Committee for Cynthia Anne Austin Certifies that this is the approved version of the following dissertation:

Investigating the Mechanisms of Therapeutic Assessment with Children: Development of the Parent Experience of Assessment Scale (PEAS)

Committee:

Deborah Tharinger, Co-Supervisor

Timothy Keith, Co-Supervisor

Stephen Finn

Stephanie Cawthon

Alissa Sherry

Investigating the Mechanisms of Therapeutic Assessment with Children: Development of the Parent Experience of Assessment Scale (PEAS)

by

Cynthia Anne Austin, B.A., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2011

Acknowledgements

I would like to express my appreciation for those whose support and contributions have made this work possible. My co-supervisors, Drs. Deborah Tharinger and Tim Keith, have provided invaluable assistance and supervision. Their mentorship has given me such a rich graduate school experience and I am so deeply appreciative of their continual guidance and support. I would like to thank the other members of my dissertation committee, Drs. Steve Finn, Stephanie Cawthon, and Alissa Sherry, for their time, suggestions, and support during this process.

This study could not have been conducted without the work of Shea Pilgrim, Jonathan Fowler, and May Matson who so graciously provided data from their research. I would also like to acknowledge the great work that is being done by the sites in this study, and the collaborative efforts with Drs. Nancy Nussbaum, April Fernando, Lisa Lasater, and Mike Brookes. For the participants in this study, thank you for your time and willingness to partake in these research efforts – your contributions made this work possible.

And to my family I would like to express my most heartfelt gratitude for your unconditional love and support. My parents, Melinda and Leon Hall, have given me both the foundation and inspiration to pursue my career, for which I will always be grateful. And for my husband, Cody, thank you for helping me make my dreams a reality. There is no one else I would rather pursue this journey with than you.

Investigating the Mechanisms of Therapeutic Assessment with Children: Development of the Parent Experience of Assessment Scale (PEAS)

Publication No._____

Cynthia Anne Austin, Ph.D.

The University of Texas at Austin, 2011

Supervisors: Deborah Tharinger and Timothy Keith

Therapeutic Assessment (TA) is a hybrid of assessment and therapy techniques in which assessors actively collaborate with clients during an individualized assessment. TA is centered around client assessment questions and provides a safe environment where clients can create shifts in their 'story' of self. More specifically, TA with children and their parents has demonstrated more confident parenting and parents' better understanding of their child's difficulties, while children have shown decreased problem behaviors and improved social/emotional functioning. The theoretical framework behind TA emphasizes the importance of the interpersonal interactions between the assessor and client, such as the development of a strong assessor client relationship and collaboration. These interpersonal processes are conceptualized as catalysts for greater depth of parent investment in the assessment and deeper levels of feedback results.

The need for greater parent involvement and partnership in child mental health services is increasingly recognized in the client/parent satisfaction literature. Parent

feedback to child mental health services is most often acquired through satisfaction questionnaires. However, the satisfaction literature has well known limitations, specifically a lack of unifying theory and methodological issues in scale development. Parent satisfaction research indicates that interpersonal experiences are more related to satisfaction than outcomes or client characteristics, and that more psychometrically sound measures are needed. Currently, satisfaction surveys do not provide a detailed understanding of parents' experiences to inform practice and research.

The current study outlines the development of the Parent Experience of Assessment Scale (PEAS). The PEAS is anchored in the theoretical orientation of TA and provides a more quantifiable measure of hypothesized underlying TA constructs. Moreover, the development of the PEAS uses advanced statistical techniques, such as Confirmatory Factor Analysis and invariance testing, to provide a higher level of psychometric rigor. The resulting scale consists of 24 items divided among 5 subscales with demonstrated relationships to general satisfaction. Structural equation modeling provides insight via direct and indirect effects among the PEAS subscales and their relationship to general satisfaction. Through the development of the PEAS, this study provides empirical evidence and support for TA theory and a more nuanced understanding of parent experiences related to satisfaction.

Table of Contents

List of Tables ix

List of Figures xi

CHAPTER I: INTRODUCTION **1**

CHAPTER II: LITERATURE REVIEW **7**

Therapeutic Assessment 8

Client Satisfaction 23

Parent Satisfaction 29

Methodological Issues 36

Psychometrics and Scale Development 47

Summary and Statement of the Problem 61

CHAPTER III: METHOD **63**

Development of the Instrument 63

Implementation 67

Research Questions and Hypotheses 69

Analysis.....76

CHAPTER IV: RESULTS **80**

Descriptive Statistics.....80

Scale Analysis and Revision 86

Additional Research Questions 105

CHAPTER V: DISCUSSION **114**

Findings for Research Questions and Hypotheses 114

Implications for Therapeutic Assessment Theory and Satisfaction 125

Directions for Future Research 131

Limitations 134

Summary.....135

Appendices.....138
A: Conceptual Diagram of Parent Interpersonal Processes in TA-C 139
B: First Order Six Factor Model of PEAS 140
C: Bi-Factor Model Confirmatory Factor Analysis for PEAS 141
D: Structural Equation Model of Direct Effects of PEAS Subscales
On General Satisfaction 142
E: First Order Base Model 143
F: First Order 27 Item Model 144
G: First Order Five Factor Model (25 Items) 145
H: Second Order Model with General Factor 146
I: Second Order PARC Best Model 147
J: Bi-Factor Model 148
K: Demographic Analysis Results 149
L: Reliability Statistics for Revised PEAS 151
M: Total Score Conversion to T-Scores 154
N: Client Satisfaction Questionnaire (CSQ-8) 156
O: Parent Experience of Assessment Scale (PEAS) 157
P: Final Version of the PEAS 161
References.....162

List of Tables

Table 1: Site Characteristics....83

Table 2: Descriptive Statistics for Selected PEAS and CSQ Items 84

Table 3: Reliability and Descriptive Statistics for Measures 85

Table 4: Fit Indices for PEAS Revision Models....88

Table 5: Fit Indices for Second Order Model Comparisons 90

Table 6: Fit Indices for Bi-Factor Model Comparisons 92

Table 7: Invariance Testing for Parent Assessor Relationship and Collaboration....96

Table 8: Invariance Testing for Systemic Awareness....97

Table 9: Invariance Testing for Negative Feelings....97

Table 10: Invariance Testing for New Understanding of Child....98

Table 11: Invariance Testing for Child Assessor Relationship....100

Table 12: Invariance Testing for Second Order Model of Revised PEAS 103

Table 13: Fit Indices for Revised PEAS 103

Table 14: Correlation of PEAS Subscales with CSQ Score 107

Table 15: Fit Indices for Revised PEAS 109

Table 16: Correlation of PEAS Subscales with CSQ Score 110

Table 17: Descriptive Statistics for PEAS Subscales by Assessment Type ...111

Table 18: Independent t-tests Results for PEAS subscale by Assessment Type 111

Table 19: Reliability and Descriptive Statistics for Revised PEAS Subscales 112

Table 20: Descriptive Statistics and Percentiles for PEAS Subscales 113

Table 21: Scoring Guidelines for Revised PEAS Subscales....113

Table 22: Analysis of Variance for Child Ethnicity.....149
Table 23: Independent t-tests for Gender.....150
Table 24: Reliability Statistics for Parent Assessor Relationship and Collaboration.....151
Table 25: Reliability Statistics for New Understanding of Child.....151
Table 26: Reliability Statistics for Child Assessor Relationship.....152
Table 27: Reliability Statistics for Systemic Awareness.....152
Table 28: Reliability Statistics for Negative Feelings.....153
Table 29: Subscale Total Score Conversion to T-Scores.....154

List of Figures

Figure 1: Revised Structural Equation Model for Parent Satisfaction 44

Figure 2: Revised PEAS Model 104

Figure 3: Structural Equation Modeling of PEAS and General Satisfaction 108

Chapter I: Introduction

Therapeutic Assessment (TA), created by Finn and colleagues, (Finn, 1996, 1997, 2003; Finn & Kamphuis, 2006; Finn & Tonsager, 1997, 2002) is a short-term intervention in which the assessor facilitates a highly collaborative 'holding environment' where clients can ask questions, explore assessment results, and create shifts in their 'story' of self. TA is designed to provide clients with a positive change experience, and can lead to increases in client motivation to continue with recommendations and services. Results from studies evaluating TA have found greater client self-esteem, increased hopefulness, greater likelihood of completing recommendations, and decreased symptomology (Finn & Tonsager, 1992, 1997; Tharinger, Finn, Wilkinson, & Schaber., 2007). More recently, TA is being explored with children (TA-C) and their families via the Therapeutic Assessment Project (TAP). Parents who have experienced TA-C report more confidence in their parenting, a better understanding of their child's difficulties, and an increase in positive affect associated with their child's future outlook (Tharinger, Finn, Gentry, et al., 2009). Children have shown decreased behavior problems and improved social/emotional functioning (Tharinger, Finn, Gentry, et al; Hamilton et al., 2009).

By drawing on both assessment and therapy techniques, TA-C provides parents the support they need become unstuck in their perceptions of their child, which can lead to new understanding and interactions between family members. Assessors in TA-C utilize a collaborative stance and involve the parents as co-investigators into their children's problems. Parents are allowed to watch their child's assessment from 'behind the mirror' (or via a live video feed) where an assessor can help parents process emerging assessment results and parent reactions 'in the moment'. The parent-assessor relationship

and collaboration experience between parents and the assessor behind the mirror is hypothesized to promote parent changes in their 'story' of their child. These changes can help parents promote a new level of understanding and compassion towards their child. In addition, TA-C fosters greater parent understanding of systemic influences and the role of context in maintaining a child's problem behaviors. Thus, parents are empowered to make changes in the child's larger systems, including their own responses, which can lead to greater hope and optimism about the child and family future.

Although the effects of TA and TA-C have been documented in previous research studies, there is not currently a measure designed to elucidate the principles underlying the parent-assessor interactions during TA-C. Specifically, the processes of collaboration, the therapeutic parent-assessor relationship, the facilitation of a new understanding of the child, and increased systemic awareness are conceptualized as interlaced experiences which serve as major catalysts in creating the positive results associated with TA-C. A parent self-report measure designed to assess the quality of these types of experiences could lend further support to the theoretical processes of TA and provide important feedback regarding parental perceptions during the child assessment process.

In the general child assessment and child mental health contexts, parent feedback is most often solicited via client or parent satisfaction surveys. Although a clear and comprehensive theory of 'satisfaction' has not been developed in the client/parent satisfaction literature, research has consistently shown that it is the interpersonal aspects, or process (Donabedian, 1988) associated with child services that is the most highly related to general satisfaction. "Research findings have consistently shown that the most important factor contributing to satisfaction in the healthcare context has been

interpersonal relationships between staff and consumers" (Gerkensmeyer & Austin, 2005, p. 62). It appears that rather than outcome variables such as symptom reduction, it is the experience of support and respect parents have while receiving services for their child which is of primary importance. According to Williams and Calnan (1991) "the specific criteria which yield the highest association with overall satisfaction scores have less to do with things such as access, availability, level and type of service provision, etc., and more to do with the nature and quality of the doctor-patient relationship and the General Practitioner's interpersonal skills" (p. 240).

The Client Satisfaction Questionnaire (CSQ-8) developed by Larsen et al. (1979) has become the gold standard in assessing general client and parent satisfaction. It is the only measure of overall satisfaction which has been used and replicated over a wide variety of studies and respondent populations. However, other researchers have focused on multidimensional aspects of client satisfaction (Greenfield & Attkisson, 2004; Essex, Fox, & Groom, 1981) that include measuring parent-practitioner interactions, such as dignified and respectful treatment.

Measuring general satisfaction with services has become a standard part of increasing consumer feedback in primary healthcare, mental health, and child mental health services. The parent satisfaction literature in particular has focused on including parent feedback as an essential part of serving children and their families. McNaughton (1994) summarizes the importance of measuring parental satisfaction because a) parents have the major responsibility and primacy in child care, b) parent satisfaction can be used to help develop program services, c) parent satisfaction may increase parent participation

by including parents in the evaluation of services, and d) parent satisfaction data can be used to convince funding audiences of the efficacy of services.

However, maximizing the potential utility of feedback from parent satisfaction data is hampered by methodological issues, including chronically high satisfaction scores. The well-known tendency for satisfaction scores to be overly positive can lead to complacency in program revision, while still maintaining the attractiveness of parent satisfaction measures as 'proof' of effective services (Stallard, 2001). In fact, researchers have consistently found that general satisfaction is not related to outcomes or client characteristics/demographic data. Thus, high scores on client satisfaction measures do not imply that the best service has been provided (Stallard) or that clients have had only positive experiences with services (Williams Coyle, & Healy, 1998)

Parent satisfaction measures are increasingly focused on the collaboration and interpersonal experiences of parents with staff and clinicians when receiving child services. For example, Summers et al. (2005) created a measure of parent partnership with special education services in educational settings. The Family Focused factor included respectful and supportive interactions between parents and staff, such as clear communication, equality, and respect for the family's values. Shifting the focus from access and logistics to the interpersonal experiences parents have with service providers indicates a need to more formally investigate the interpersonal processes repeatedly associated with the highest levels of satisfaction.

Perhaps the greatest difficulty faced by the client/parent satisfaction literature has been the creation of dozens of satisfaction measures designed for site specific use without appropriate scale revision techniques. Most notable has been the default use of Principal

Components Analysis (PCA) when factor analysis provides better theoretical and statistical support in scale development. Factor analysis investigates the underlying psychological construct or latent variable of interest by accounting for common and unique variance (including error) among item covariances. Most psychological constructs that researchers want to measure are latent variables that can only be inferred through direct performance on a test or response to a survey item (DeVellis, 2003). PCA does not account for the unique variance of items and therefore, despite similar results, is not theoretically or statistically warranted in most scale development.

Confirmatory factor analysis (CFA) requires researchers to have a strong a priori theoretical model of factors and pattern of factor loadings. CFA then allows researchers to test hypotheses of the fit between a specified model and the data, and test competing models to increase statistical support or make informed revisions for a model. Most researchers in the satisfaction literature implicitly assume that significant differences in scores on satisfaction measures represent a significant difference in the latent variable of 'general satisfaction'. However, this assumption has not been tested, despite the availability of multiple group mean and covariance structure analysis (MG-MACS). MG-MACS uses confirmatory factor analysis to test for invariance across items to ensure that items are functioning similarly across groups. This type of higher level analysis is common in test development, but has not been used within the satisfaction literature. Incorporating statistical techniques such as CFA and MG-MACS could provide greater credibility and evidence of quality scale development for a parent self-report measure.

Overall, it appears that TA-C is already putting into practice principles that the parent satisfaction literature is investigating, such as parental collaboration and the

relationship between parents and clinician. The present study traces the development of a theoretically based and methodologically sound parent self-report measure, which will investigate the interpersonal relationships and experiences of parents that are hypothesized to be highly related to general parental satisfaction. CFA and MG-MACS will be used to heighten the psychometric properties of the measure and the overall level of rigor in the scale development. By creating a parent measure that explores the interpersonal experiences of parents during a child assessment, the current study will simultaneously investigate the processes most related to parent satisfaction and provide more verification of the collaborative constructs of TA.

Chapter II: Literature Review

The following literature review begins with an overview of Therapeutic Assessment (TA). It then proceeds to the client and parent satisfaction literature base in assessing general and multidimensional conceptualizations of satisfaction related to child mental health services. The development of TA and its efficacy and effectiveness with both adult and child clients is described. Following this overview is an elaboration on the types of collaborative parental interactions that occur between the parents and the assessor in TA with children. The assessor-parent relationship and active collaboration with parents developed over the course of the child assessment are conceptualized as the key catalysts to increases in parental openness to new information, re-investment in the child, greater awareness of family involvement, and increased follow through on recommendations.

Parent satisfaction is a piece of the client satisfaction literature spanning the last 30 years. Research in this area has found repeated evidence that interpersonal experiences and relationships, such as those in TA, are more associated with general satisfaction than outcomes, demographics, or client characteristics. The analysis will review the most prevalent methodological issues throughout the client satisfaction literature including questionnaire development, psychometric properties, and sampling methods, and will conclude with a discussion of the appropriate use of confirmatory factor analysis and testing of measurement invariance in scale development.

Therapeutic Assessment

Overview

Therapeutic Assessment is a form of collaborative assessment that is individualized based on client questions and designed to be a "collaborative, short term intervention" (Tharinger, Krumholz, Austin & Matson, in press). TA has been developed by Finn and colleagues (Finn, 1996, 1997, 2003; Finn & Kamphuis, 2006; Finn & Tonsager, 1992, 1997, 2002) and has been reciprocally influenced by the development of collaborative assessment (Fischer 1985/1994; Handler 1996, 2007). As summarized by Finn and Tonsager (1997), TA has its roots in the humanistic movement by advancing the idea of sharing test results with clients. This more collaborative approach allows for a greater therapeutic alliance and the setting of specific goals for the assessment. Early studies (Finn & Tonsager, 1992) found an increase in client self-esteem, motivation, and feelings of hope, while reducing client symptomology and feelings of isolation. TA is explicitly designed to do more than provide a diagnosis and recommendations. It is conceptualized as a "semi-structured hybrid of assessment and intervention techniques" that challenges assessors to employ both assessment and therapeutic skills in providing a transformative experience for clients (Tharinger, Krumholz, et al.; Finn, 2007).

TA can be contrasted with the more typical information-gathering model of assessment where the assessor is an objective observer gathering data from the client. In TA, the assessor actively enlists the help of the client to question, explore, and test assessment information. This allows the client to learn new ways of thinking about self and others in a supportive environment that can lead to lasting changes in the client's story of self (Finn & Tonsager, 1997, Tharinger, Finn, Wilkinson, & Schaber, 2007).

However, it is important to note that TA does not require assessors to sacrifice obtaining valid and reliable test information. Rather, TA can provide additional conceptualizations of the client's assessment experience, by processing client responses, affect, and strategies after the completion of standardized measures.

The principles of TA allow the assessor to go beyond the role of the impassive assessment expert by using test responses and experiences to "get in the shoes" of clients for empathic understanding (Finn, 2007). The assessor's own reactions, affect, and counter transference are also important pieces of information used to reveal case dynamics and potential sources of bias (Tharinger, et al., 2007). One of the main techniques used in TA is 'collaborative empiricism' as found in cognitive behavioral therapy (Tharinger et al., 2007) where the assessor and client co-investigate various experiments about the client's questions and expected results. By engaging the client in the conceptualization and interpretation of these experiments, the assessor can help guide the client through the assimilation of new information and experiences (Finn & Tonsager, 1997, 2002; Tharinger et al., 2007).

Research with TA has shown positive treatment effects for adults, including increased likelihood of completing recommended treatment, decreased symptomology, greater self-esteem, and increased hopefulness (Ackerman, Hilsenroth, Baity, & Blagys, 2000; Finn & Tonsager 1992; Tharinger et al., 2007). Case studies, including work with adults and children/adolescents, have shown the clinical utility of TA and the effects on parents (Hamilton et al., 2009; Tharinger, Finn, Gentry, et al., 2009). Parents have reported gaining a better understanding of their child's problems, and feeling more confident in their parenting and in seeking additional services, whereas children have

shown decreased behavior problems and improved social/mood functioning (Tharinger et al., 2007). Parents have also shown increased positive affect and decreased negative affect relating to their child's challenges and future outlook (Tharinger, Finn, Gentry, et al.). According to the evidence of 14 case studies from the Therapeutic Assessment Project (described below), parents reported a significant decrease in child symptomology and mothers and children reported enhanced family functioning. Mothers demonstrated significant increases in positive emotion, significant decreases in negative emotionality concerning the child's challenges and future outlook (Tharinger, Finn, Gentry, et al.; Tharinger, Krumholz, et al., in press).

Client Experiences with Therapeutic Assessment

Previous work has been conducted in developing a measure of adult client experiences with assessment by Finn, Schroeder, and Tonsager (2004). The Assessment Questionnaire - 2 (AQ-2) was based on an earlier 30 item version created by Finn and Tonsager "to measure the subjective reactions of clients from a university counseling center to a brief assessment" (p. 4). Using a sample of one hundred twenty three college students, the fifty six items of the AQ-2 were subjected to principal component analysis with oblique rotation. The results found a four factor model translating into the following subscales: New Self Awareness/Understanding ($\alpha = .89$), Positive Accurate Mirroring ($\alpha = .87$), Positive Relationship with Assessor ($\alpha = .87$), and Negative Feelings About the Assessment ($\alpha = .85$).

Due to the significant correlations among subscales, a test for a higher-order factor was conducted by using a principal factor analysis on the correlation matrix. A one-factor solution was superior and accounted for 70% of the variance among the four

subscales. The higher-order factor, named Positive Experience, appeared to differentiate among overall positive versus negative evaluation of the assessment experience. Lastly, discriminate validity was provided by correlations below r = .05 of scores on the AQ-2 subscales and Positive Experience factor with scores on the Marlow-Crowne Social Desirability Index (Crowne & Marlow, 1961).

A second study was conducted to determine reliability of the AQ-2 with differing populations and to investigate test-retest reliability. The three samples included a) college low self-esteem participants (n = 73), b) inpatient sample (n = 35), and c) outpatient sample (n = 41). Each subscale demonstrated acceptable to excellent reliability coefficients across the three samples: New Self-Awareness ranged from .84 to .93; Positive Mirroring from .88 to .89; Positive Relationship from .89 to .90; Negative Feelings from .79 to .92. The test-retest sample, using a two-week interval from the college sample, ranged from .75 to .84 for the AQ-2 subscales.

The development of the AQ-2 provides a preliminary framework for providing empirical evidence for client's reactions to psychological experience. The four subscales on the AQ-2 support a multidimensional process for clients including gaining new information about themselves, confirmation of positive self-attributes from the assessor, having a mutually positive relationship with the assessor, or feeling negative about the assessment due to feeling judged or uncomfortable during the assessment (Finn, Schroeder, and Tonsager, 2004). The authors conclude that the general factor of Positive Experience indicates that "client's positive experiences with assessment are greatly dependent on the empathy and positive regard demonstrated by assessors" (p. 12). The AQ-2 was created based on the framework of TA in an effort to "allow psychologists to

better tailor their assessment services to meet client needs" (p. 12). The current study will examine similar constructs, but in the context of seeking the reactions of parents based on Therapeutic Assessment as applied with children.

Therapeutic Assessment with Children (TA-C)

The Therapeutic Assessment Project (TAP) is a systematic research study investigating TA with adolescent and preadolescent children and their families (Tharinger, Finn, Gentry, et al., 2009). When working with children and their families, TA becomes a short family systems intervention whose goal is to help families become 'unstuck' concerning the identified child (Tharinger et al., 2007). Parents are key collaborators in the assessment process who are guided to a more empathic understanding of their child and who learn ways to shift family interactions toward more positive outcomes. Through the 14 case studies completed by TAP, researchers have attempted to extrapolate the underlying mechanisms and processes that make TA with children a success. Tharinger et al. (2007) outlines the general structure of the assessment process including type and order of the typical 8-10 sessions with a detailed case example. The six phases of TA conceptualized by Finn (2007) include the a) assessment question gathering phase b) standardized testing phase, c) family intervention phase, d) summary discussion phase, e) written communication phase, and f) the follow-up phase. The use of the full TA model exemplified in TAP can be considered 'comprehensive TA-C' as described in Tharinger, Krumholz, et al. (in press) and Tharinger, et al. (2007). The following are some of the unique aspects of TA-C that highlight the collaborative process, such as helping parents make shifts in their understanding of their child and in

their own role in helping their child and family cope with mental health and behavioral struggles.

Assessment Question Gathering Phase. The first phase in TA is to solicit client questions about the purpose of the assessment and help the client formulate questions that can be investigated via the different techniques available for TA (Finn, 2007). In TA-C, the parents meet with the assessors to give initial questions and background, and in the subsequent session the assessor helps the parents share a developmentally appropriate question with the child. Children are then asked to contribute their own questions, both with the parents present and in the following one-on-one activity time with the assessor (Tharinger et al., 2007).

The question-gathering process provides the structure for the ensuing assessment sessions and allows the assessors to choose the most relevant testing methods such as interview, observation, psychoeducational tests, neuropsychological tests, self reports, behavior rating scales, and performance based personality measures (e.g. drawings, sentence completions, TAT, Rorschach). An advantage of the TA methodology is that standardized test scores can be collected to provide a nomothetic perspective (Finn & Tonsager, 1992; Tharinger et al., 2007). However, extended inquiry, testing of the limits, and processing of the assessment experience allows for an individualized assessment that may better address the concerns and questions the parents/child bring to the assessment. As will be discussed in a subsequent section on parent satisfaction, investigating parent expectations is often an overlooked aspect of child mental health services. By engaging parents in the question gathering process, TA allows assessors to learn and provide feedback to parents about what an assessment can realistically offer.

The assessment questions also connect the assessment to the parents' practical, daily concerns about their child, and create an opportunity for increased parent investment and participation in the assessment process. The questions can be revised throughout the assessment and serve as the anchor for the discussion feedback sessions at the end of TA-C (summary/discussion phase). It takes a conscious and purposeful effort by the assessor to engage the clients in question making, without falling into the typical expert role or accepting vague or unrealistic questions. This process sets the stage for the development of the parent-assessor and child-assessor relationships that will develop over the course of the assessment and are vital to successful collaboration.

Standardized Testing Phase. One of the most revolutionary aspects of TA with children is that during the standardized testing phase, parents are encouraged to watch their child's assessment sessions (either in the room, through a one way mirror, or via a live video feed) (Hamilton, et al., 2009; Tharinger et al., 2007; Tharinger, Krumholz, et al., in press). This 'behind the mirror' technique allows one member of the assessment team to help inform parents about general testing procedures, encourage questions and reactions to the child's performance/behavior/affect, and guide parents to gaining a new perspective of their child. The child is aware of the parents behind the mirror, and often uses this instrument to communicate with parents. In follow-up interviews of a TA-C, parents often cite the 'behind the mirror' to be one of the major benefits to seeing their child in a new way and watching their child interact with different people (Hamilton, et al.; Tharinger., Finn, Arora, et al., in preparation). Current research on TAP is focusing on the collaboration behind the mirror between the parents and assessment team to

further highlight how this process enhances change in parent perceptions and learning. A more thorough discussion of these processes follows the general overview of TA-C.

Family Intervention Phase. The family intervention phase is an opportunity for the parents to be directly involved in the child assessment. After gathering assessment data, the assessment team devises an activity for family members to engage in ranging from simply playing a game, to a consensus Rorschach, to re-enacting a family argument. The family session has multiple goals including bringing the assessment findings to life, testing more systemic hypothesis of the child's problems, and allowing the family to have a positive experience to generate new ways of interacting (Tharinger, Finn, Austin, et al.). The family session can often serve as a major 'ah ha!' moment where parents can see how their reactions influence their child and gain new skills. The variety of techniques available (play therapy, empathic listening, psychodrama, consensus TAT, etc.) allows the assessment team to craft an individualized family experience designed to meet the family at their current level of understanding and then hopefully create the catalyst for further assimilation of the assessment results (Tharinger, Finn, Austin, et al.). The family session is a prime example of not only having parents highly involved in a child assessment, but also parents being actual participants just as they would need to participate in successful treatment or therapy strategies.

Summary Discussion Phase. One of the earliest findings of research in TA was that feedback about testing is more readily assimilated and understood if ordered according to client's preconceived perceptions (Finn 1996; Finn & Tonsager, 1997). This has led to the recommendation of presenting client feedback from the most congruent to increasingly more discrepant order of client understanding. In TA-C, there are unique

aspects when considering how to give developmentally appropriate child feedback and how to order feedback to be of the most use for parents (Tharinger, Finn, Hersch, et al., 2008). When providing feedback to parents in TA, it is divided into three levels, with the assessment team judging by the parents' reactions how to proceed to increasingly more difficult information. However, the case studies on TAP have found that since parents are continuously receiving feedback and incorporating new perspectives as an ongoing process, by the end of the assessment parents are able to hear more systemic pieces of feedback at the higher levels (Hamilton, et al., 2009; Tharinger, Finn, Hersch, et al.). This assimilation is also aided by the feedback information being organized around the parent and child assessment questions, thus anchoring the results around areas already personally relevant to the clients.

Written Communication Phase. In continuing with the collaborative nature of the assessment, parents are invited to make sure the results 'fit' with their experience and to add suggestions or their own interpretations of the assessment experience. After the summary/discussion session, parents are also given a letter of the assessment results organized around their assessment questions. Formal reports are also provided as needed for schools and other services. The child feedback is often presented as a fable crafted by the assessment team (and approved by the parents) that speaks metaphorically to the most salient struggles the child is facing and how the parents will be able to help the child cope (Tharinger, Finn, Hersh, et al., 2008; Tharinger, Finn, Wilkinson, et al., 2008). By organizing and presenting feedback using the collaborative techniques (levels, fable writing, letter writing, and parent discussion session) cited in the TAP articles, TA with

children is demonstrating how assessors and professionals can communicate in a way that parents understand and that keeps them involved in the entire process.

This overview has provided a brief synopsis of TA-C in its comprehensive form. However, it is clear that not all assessment practitioners will have the time or resources to implement the comprehensive model. Currently, studies are under way to investigate how aspects of TA with children (question gathering, feedback levels, child fables, etc.) can be transferred to school, community, and private assessment environments to demonstrate the effectiveness of TA concepts and techniques in augmenting traditional assessment practices. Tharinger, Krumholz, et al., (in press) provide a heuristic for incorporating TA methods into a school setting and case examples using collaborative/TA assessment techniques in assessments for special education. Other studies have focused on using TA to empower adolescents receiving special education assessments or incorporating a child feedback session and personal fable in child neuropsychological assessments. Positive results from these investigations are spurring more studies designed to investigate the efficacy of TA practices and techniques.

Interpersonal Processes with Parents

The ability for one member of the assessment team to process information with parents 'in the moment' behind the mirror (or watching a video feed) "significantly advances the collaborative experiences of the parents and also helps them to digest the information their child is providing through tests and creative methods" (Tharinger, Krumholz, et al., in press, p. 9). In addition to brief 'check-ins' with parents before or after the standardized assessment sessions, the time spent behind the mirror provides the

most direct opportunity for an assessor to collaboratively engage with parents in processing the child assessment results and experience.

The various tasks of the assessor working behind the mirror have been conceptualized into 13 functions (Tharinger, Finn, Arora, et al., in preparation) that reflect the various interactions between the assessor and the parents. These tasks will serve as examples of some of the general processes that are hypothesized to serve as catalysts for parental engagement and change in TA-C. The categories of processes include: parent-assessor relationship, collaboration, new understanding of child, and systemic awareness. All of the tasks described function as an iterative, circular process between the assessor and parent in experiencing the child assessment, and thus each informs and promotes the effectiveness of the others. See Appendix A for conceptual diagram.

The parent-assessor relationship is conceptualized as the trusting and respectful relationship actively nurtured by the assessor throughout TA. The importance of rapport is generally mentioned in assessment practice to help enable good motivation and hence, valid results. However, in TA, the parent-assessor relationship that develops is the supportive and emotionally rich 'holding environment' of a client-therapist relationship. The assessor strives to create a climate of warmth, acceptance, and unconditional positive regard (Tharinger, Krumholz, et al., in press). The tasks in this area include a) fostering trust between parents and the assessment team, b) emotionally supporting parents as they reach new understandings or are confirmed in their existing understandings, and c) continually assessing parental readiness for change. The assessor wants to optimize a balance between helping parents discover and embrace new conceptualizations, yet

ensuring the assessor does not push too early or too much, thereby resulting in raising parents' defenses. It is this supportive, therapeutic, relationship that is hypothesized to help parents consider and accept new information in a non threatening manner. By supporting parents emotionally throughout the assessment, TA posits that parents are able to more readily assimilate especially difficult or potentially overwhelming assessment findings.

Collaboration is a key component in all aspects of TA, and as mentioned previously, begins the first session by including the parents as co-investigators who come up with personally relevant assessment questions. Collaboration includes the assessor's conceptualization of working with the parents as part of a team, which then leads to techniques which help facilitate parents into active, invested, collaborators. The tasks in this area include a) educating parents about psychological tests and other assessment procedures, b) gathering relevant background information about the child/family, c) consensually validating certain perceptions, and d) gently confronting other perceptions by asking parents to note data which conflicts with the existing story and/or by respectfully offering different interpretations of events. All of these tasks involve engaging the parents as informed, active participants in the understanding and conceptualization of the child assessment process and results. Although the assessor remains the 'expert' on assessment, the parents are considered 'the experts' on their child, and thus a collaborative stance taken by the assessor is essential to foster participation, questioning, learning, and eventually shifts in parent understanding of their child.

Given the foundation of the parent-assessor relationship and high levels of collaboration, the assessor assists the parents in exploring new ideas and constructing a different 'story' for their child. The new understanding of child category entails fostering new conceptualizations for the parents both in terms of knowledge about their child and in terms of compassion and acceptance relating to their child's difficulties. The tasks designed to facilitate this process include a) fostering parents' curiosity about their child and the assessment process, b) gathering information about how parents perceive their child, c) helping parents notice similarities and differences in the child's behavior in the problem situation and assessment situation and d) helping them 'step back' and look with new eyes.

It is important to note that regardless of the category in which each task is placed, all tasks can be approached in a collaborative manner. For example, "gently confronting other perceptions by asking parents to note data which conflicts with the existing story and/or by respectfully offering different interpretations of events" could also be placed within the new understanding of child category, as it is clearly attempting to shift parents' conceptualizations of their child. The descriptors of 'gently confronting' and 'respectfully offering' exemplify the collaborative manner of the assessor, while the more concrete task of asking parents to note conflicting data and offering different interpretations is part of the new understanding of child process. However, the collaborative mindset of this approach is essential to successfully accomplishing this task, and its depiction of offering alternatives in a non-threatening, tentative way seemed to exemplify the collaborative approach taken throughout the assessment, thus its primary categorization under the collaboration category.

The last category, systemic awareness, consists of facilitating an even greater or deeper level of understanding by parents about what may be contributing to their child's difficulties. These tasks include a) modeling psychological mindedness and 'looking below the surface', b) observing (and processing) parents' reactions and their interactions with each other, and c) helping parents think about contextual influences on behavior. Although a child's problems are now generally conceptualized by professionals to include the family, school, and larger community systems, often treatment still centers on the individual child. Traditional assessment practice tacitly supports this focus on the child by not involving parents or other systems in the assessment process. (Tharinger, Finn, Austin et al., 2008). The active and collaborative role of parents in the TA-C process is designed to promote greater awareness of systemic or family influences that may be contributing to or maintaining a child's difficulties. This increased systemic awareness is also used to help empower parents to re-engage as positive influences in the family system and engender hope. Helping parents to gain this type of awareness is one of the strongest and unique aspects of TA-C, and is created by incorporating 'mini feedback' and hypotheses concerning parent and family processes throughout the assessment, in particular through comments and questions made by the assessor behind-the-mirror. The behind the mirror tasks help prepare parents for this new awareness long before the summary/discussion meeting and can help the assessor form hypotheses for the family session phase discussed previously.

Overall, evidence suggests that adult and child assessment using the principles and techniques of TA function as a direct, positive intervention. Rather than helping clients indirectly through recommendations given at the end of the assessment (which

may or may not be implemented), TA posits that assessment in and of itself can be a transformative experience (Finn, 2007; Tharinger, Krumholz, et al., in press) To date, the effects of TA-C have been measured via quantitative analysis, pre/post interviews with parents and child, various weekly session self reports (alliance with assessment team, family alliance, positive/negative affect), pre and post BASC completion, and 6-month follow-up measures (Tharinger, Finn, Gentry, et al., 2009). The work on the Assessment Questionnaire – 2 has indicated that client reactions to assessments are multidimensional, but strongly influenced by the overall positive relationship and support from the assessor. There is not currently an established self-report instrument designed to measure the specific interpersonal aspects hypothesized to be present for parents during TA with children. It is these interpersonal experiences, such as the parent-assessor relationship, collaboration, and child-assessor relationship, that are conceptualized as the necessary catalysts for a successful TA.

A measure designed to quantify these experiences would potentially provide support for the TA process and help distinguish the unique aspects of TA when compared to traditional assessment modalities. For instance, a parent self report (and eventually, a child self report) highlighting these constructs could show a significant difference between a traditional information gathering assessment with low levels of parent collaboration as compared to TA-C with high levels of parent input. “Whenever possible, we need to utilize methods of assessment and intervention that are responsive to consumers and have proven efficacious” (Tharinger, Krumholz, et al., in press, p. 1). A measure of parent experiences of a child assessment could help provide more empirical evidence for underlying TA concepts and the importance of collaborative interpersonal

interactions in working with parents during a child assessment. In particular, because TA-C is conceptualized as a family intervention, it is important to gather more data on exactly which aspects of the TA experience are most salient to parents in promoting a positive experience and change in the 'story' of their child. TA has much to offer the field of child assessment and as the proliferation of TA principles and techniques continues, a way to validate the underlying processes becomes essential in promoting further research and investigation across expanding contexts.

Currently, the most common method for collecting parent feedback from a child assessment, or child mental health services in general, is via a client/parent satisfaction questionnaire. Thus, in order to create a parent measure with the most applicability to various sites and programs of child assessment, a review of salient aspects of client and parent satisfaction is warranted. Understanding research findings and methodolgocial/scale development issues within the client/parent satisfaction literature can help inform the creation of a sound measure of parent experiences and interpersonal processes during their child's assessment.

Client Satisfaction

Background and General Findings

For the last 30 years, client satisfaction measures have become a routine part of consumer feedback collected by Community Mental Health Centers (CMHC) (Essex, Fox, & Groom, 1981; Harrington Godley, Fiedler, & Funk, 1998; Lebow, 1982). The movement has been the result of increased pressure on CMHC's to provide accountability in an increasingly consumer-oriented society (Essex, et al.; Lebow; Plante, Couchman, & Hoffman, 1998). The early progress of the client satisfaction movement is adeptly

summarized in reviews by Lebow and Larsen, Attkisson, Hargreaves, and Nguyen (1979) which discuss the state of client satisfaction literature once consumer feedback (most often via surveys) had become "a standard part of the practice of many mental health facilities" (Lebow, p. 244). Initially, the inclusion of client feedback as a valid perspective had to be supported by developing research, fostered in part by legislative mandates such as the Community Mental Health Centers Amendments of 1975. This mandate required the measuring of "acceptability of practices to the client/patient" (Larsen et al., 1979). An increased emphasis on accountability in Community Health Centers also resulted in including the client in the evaluation of services, most often through client satisfaction measures (Essex, et al.). Larsen et al. argue that "when the client's perspective is not taken into account, the evaluation of services is incomplete and biased towards the provider's or evaluator's perspective" (p. 197).

Many of the conclusions of Lebow (1982) and Larsen et al., (1979) are still supported today, namely that demographics have not been found to be good predictors of satisfaction (Essex, et al., 1981; Harrington et al., 1998; Lebow; Measelle, Weinstein, & Martinez, 1998; Young, Nicholson, & Davis, 1995), that satisfaction ratings are generally high, between 70-80%, (Riley, Stromberg, & Clark, 2005) or "uniformly positive" (Essex et al., p. 227), and that other outcome measures, including therapist satisfaction, therapist ratings of client satisfaction, and client rated outcome measures lack consistent results (Larsen et al.; Lebow). Despite their widespread use, reviews of client satisfaction literature have consistently found a lack of standardization, psychometric validation, and resolution of methodological issues in satisfaction measures (Lebow; Young, et al.). Many satisfaction surveys are created in-house by the mental health centers without the

background knowledge or financial resources to develop comprehensive outcome/satisfaction or program evaluation measures (Plante et al., 1998).

Although the client satisfaction literature has focused primarily on mental health settings, there is an equally large literature base in patient satisfaction in primary healthcare settings. Pascoe (1984), in his review of patient satisfaction, describes satisfaction as the "recipient's reaction to the context, process, and results of their service experience" (p. 189). Donabedian (1988) defined satisfaction in terms of three aspects: structure, process, and outcome. Structure refers to the attributes of the setting, process is what occurs in the actual giving and receiving of services, and outcomes are the direct effects of care on client. A meta-analysis listed the top 12 elements of satisfaction, and the most often researched aspect was Humaneness (warmth, respect, kindness, willingness to listen, appropriate non-verbal behavior, and interpersonal skill) included in 30% of studies. According to Williams and Calnan (1991) "the specific criteria which yield the highest association with overall satisfaction scores have less to do with things such as access, availability, level and type of service provision, etc., and more to do with the nature and quality of the doctor-patient relationship and the General Practitioner's interpersonal skills" (p. 240). Lewis' (1994) review of patient satisfaction with general care across 41 studies found satisfaction was most related to amount of information given, greater technical and interpersonal competence, more partnership building, more social conversation, more positive and less negative talk, and more overall communication.

Thus, by the early 1990s, the patient satisfaction literature had recognized the importance of the interpersonal process as related to general satisfaction. "From a

theoretical point of view, the evidence supports the arguments of those who have highlighted the significance of the social and psychological elements in the therapeutic relationships between health professionals and their clients" (Williams & Calnan, 1991, p. 715). Lewis (1994) described these interpersonal processes as 'bedside manner' or the 'art of care' and recognized that, though not new, the importance of these processes may be "periodically overlooked" but they "appear to be the aspect of care most consistently demanded by the consumer" (p. 668).

Sheppard (1993) also emphasized the importance of interpersonal skills for practitioners in the client satisfaction research by citing the need for a "dialogue of communication, empathy, the openness of the clinician, and client participation in planning/intervention" (p. 257). According to Greenfield and Attkisson (2004) "the satisfaction construct occupies a position centrally related to the 'working alliance' that exists between service provider and client" (p. 820). The development of a multifactor scale of client satisfaction, the Service Satisfaction Scale (SSS) has a primary healthcare version (SSS-15) and a mental health version (SSS-30). Both versions contain two factors, a) Practitioner Manner and Skill and b) Perceived Outcome (Greenfield & Attkisson, 1989). The scale is designed to assess interpersonal manner, technical quality, efficacy/outcome, accessibility/convenience, physical environment, and availability. However, the domains associated with 'structure' such as accessibility, physical environment, and availability were less consistent, and the developers suggested those domains may be more service setting-specific than the other two general factors (Greenfield & Attkisson, 2004).

Other researchers in the client satisfaction literature found additional evidence for a multidimensional view of satisfaction and the importance of interpersonal constructs. Essex et al., (1981) developed a questionnaire via factor analysis with four dimensions: Satisfaction with Services, Acceptability of Clinician, Impact of Services, and Dignified Treatment. Overall satisfaction in and of itself did not include the client perceptions of effectiveness of the service (agreement on goals, right type of service, recommend to others), acceptability of the clinician (age, race, sex) and client treatment (dignity, respect, confidentiality, promptness, agreement on termination). Essex et al. suggested that their findings supported earlier work that satisfaction does not directly equal success (or symptom reduction) and hence client perceptions of various dimensions are necessary to gain a full picture of a client's experience with services.

Client Satisfaction Questionnaire

Despite the importance of a multidimensional approach to satisfaction and the influence of interpersonal processes, most researchers and programs either create or use a measure of 'general' satisfaction. The development of the Client Satisfaction Questionnaires (CSQ) by Larsen et al., (1979) was a response to researchers struggling "to construct a psychometrically adequate scale with demonstrated validity, brevity, low cost, and ease of administration" (Attkisson & Zwick 1982; p. 233). A major issue with early and continuing client satisfaction research is that it is often locally generated and consumed (Larsen et al., 1979; Lebow, 1982) resulting in surveys and studies that are not standardized or easily compared across settings and services (Plante, et al., 1998; Riley et al. 2005). The CSQ is now the most widely used measure for general satisfaction, and the only well-standardized adult measure (Attkisson & Zwick; Gaston & Sabourin, 1992;

Larsen et al.) now being used for parent satisfaction (Bodin et al., 2007; Byalin, 1993; Garland, Haine, & Boxmeyer, 2007; Gerkensmeyer & Austin, 2005; Harrington Godley et al., 1998; King, Cathers, King, & Rosenbaum, 2001; Plante et al.).

The original CSQ was developed via a literature search for items/concepts, 32 judge rankings of how well items tapped dimensions, and analysis of the initial 248 response sample via principal components analysis, which resulted in a single dimension with a coefficient alpha of .93. The CSQ-8 (8 item general scale) uses a 4-point Likert scale and has been found to be the shortest and most robust version for measuring general satisfaction (Attkisson & Zwick, 1982). A review of the CSQ-8 by Attkisson and Greenfield (2004) provides evidence of its wide spread use; the CSQ-8 has been translated into 15 languages including Chinese, Dutch, English, French, Japanese, Russia, and Spanish. The CSQ-8 has demonstrated construct validity with high correlations (r = .6 to.8) with other satisfaction measures (Attkisson & Greenfield). Other researchers report coefficient alphas for the CSQ-8 between .93 and .96 (Byalin, 1993; Gerkensmeyer & Austin, 2005; Plante et al., 1998). Attkisson and Greenfield provide a table of the psychometric properties of the CSQ-8, with the majority of samples from mental health facilities. The average scores range from 26.35 to 27.80 (out of a possible 32 points) with standard deviations from 3.57 to 4.30.

Initial concerns about using a measure standardized on adult clients with the parents of children receiving mental health services (Young et al., 1995) have been addressed with the findings that parent responses are often similar to those of adult clients (Essex et al., 1981) and the high coefficient alphas found with parent populations (Byalin, 1993; Gerkensmeyer & Austin, 2005). Thus it appears that the CSQ-8 is a measure of

general satisfaction that is well standardized via repeated use for both adult client and parent responses.

Parent Satisfaction

The shift in research to parent satisfaction with child mental health services became more prominent in the 1990s and has followed a similar developmental path of the original client satisfaction literature, including standard problems with methodology (see below). Not only should clients be able to provide feedback about services, but parents should also "be an integral part of the treatment of their children" (Young et al., 1995, p. 220) in all areas, including evaluation, planning, and implementation. Recent research has shown that child and adolescent mental health is best treated by meeting family needs (Riley et al., 2005) and evaluating the 'bundled' services (psychotherapy, group therapy, case management, parent support groups, social skills, etc) which have become more common service modalities (Harrington, Godley et al., 1998).

Just as the validity of including client perspectives in evaluation had to be established in the client satisfaction literature (Larsen et al., 1979; Lebow, 1982), parent perspectives were not welcomed unquestionably into the program evaluation literature on child and adolescent services. Perhaps even more so than adult clients, parents have historically been blamed for their child's problems (Measelle et al., 1998; Young et al., 1995). This perception has often been overlooked in the parent satisfaction literature (Young et al.). Thus, parents may be skeptical of service providers and inclined to under-utilize services, or reject services when dissatisfied (Measelle et al.). Similar to adult client satisfaction research, dissatisfaction with the patient-physician relationship can

predict poor treatment, under utilization of services, and premature termination (Attkisson & Zwick, 1982; Measelle et al.).

Parent satisfaction research also parallels the client satisfaction area with the findings that demographics (child gender, child age, child race, child grade, length of treatment, parent's age, parent's gender, parent education, employment status, parent income, marital status, parent race) are not significantly related to parent satisfaction (Gerkensmeyer & Austin, 2005; Harrington Godley et al., 1998; Measelle et al., 1998; Young et al., 1995). Rather, only severity of child illness and differential settings (child living at home vs. not, state vs. community settings, public vs. alternative school) have been found to predict parent satisfaction (Gerkensmeyer & Austin; Harrington Godley et al.). Rey, O'Brien, and Walter (2002) found that dissatisfied parents received more welfare services and their children had more disruptive disorders. King et al. (2001) found that as the number of child problems increased, there was a decrease in parent satisfaction. These findings appear to be related, in that children with more severe problems are more likely to be in a restricted setting, such as alternative school or state hospital. Thus, lower levels of satisfaction are associated with families who face more severe child problems that will be harder to successfully treat.

Combining the severity findings above with the findings of Plante et al., (1998) that parents report high levels of satisfaction, despite a lack of symptom reduction from treatment, indicates that "care" (child and parent support) is more important than "cure" (p. 54). Rey, et al., (2002) reported small or insignificant relationships between parent satisfaction and child improvement, citing that quality of care may be more important than its effectiveness. Riley et al. (2005) took parent responses for five factors and found

the percent of respondents with an average score greater than 3.5 on a Likert scale from 1 to 5. For the factors of Cultural Sensitivity, Access to Services, Parent Participation, and Appropriateness 70-82% of parents had an average factor score above 3.5, whereas only 47% of parents had average ratings above 3.5 for the Outcome factor. Thus, the Outcome factor had the lowest ratings and was the least associated with parent satisfaction. Just as with chronic medical conditions, such as diabetes, most childhood mental health issues (ADHD, Bipolar Disorder, Depression, Anxiety, Autism) will not disappear with treatment; rather, the goal of child and family services is to manage the symptoms successfully and try to minimize the social, emotional, and developmental side effects over time (co-morbid disorders, delinquency, gang membership, etc). King et al. (2001) notes that there may be differences for children with chronic conditions who require ongoing service delivery throughout their lives. The need for family support when dealing with child mental illness may help explain why parental satisfaction is not highly correlated with outcome measures and why it is essential for professionals to make a more formal and conscientious effort to include parents in child treatment.

More recent research is recognizing that although the child may need services, the parent is a critical component in "engagement and continuation of treatment" (Martin, Peter, & Kapp, 2003; Riley et al., 2005, p. 88) as children are dependent upon adults and do not seek services themselves (Young et al., 1995). Gerkensmeyer and Austin (2005) succinctly summarized the major role of parents who: a) obtain services for the child, b) are a key to child success via parent participation, c) are the best source of information about the effects of caring for a child with mental health problems, and d) are the primary caregivers of children after the completion of services (p. 61). The shift to including

parent perspectives in treatment evaluation, and also including parents in planning and implementation of child services requires that researchers establish the process variables associated with child treatment that are most related to parent satisfaction and a positive overall parent experience.

It is important to note that the engagement of parents in child services can also include adoptive parents, foster parents, and guardians who may seek or assist with a child mental health services. Robertson (2006) discusses the importance of including these parenting caregivers in child assessments given the "interdependence of assessment and intervention", particularly with young children who are at greater risk for developing psychological and behavioral difficulties (p. 1). Although including parenting caregivers in child assessment processes may be seen as asking even more of a newly stressed system, "when foster parents are engaged in the full process of providing assessment and services to their foster child, evidence suggests that the outcomes for foster children improve" (p. 187). When evaluating the Comprehensive Family Services (CFS) paradigm, Huebner, Jones, Miller, Custer, and Critchfield (2006) investigated the satisfaction of parents as well as foster/pre-adoptive parents. The CFS approach emphasizes family-centered services and strengths based practices, such as including clients as active participants in decision making and intervention planning. The researchers found statistically significant differences in client satisfaction, for both parents and parenting caregivers, for those provided the more collaborative CFS services. Thus, including parenting caregivers in research on parent satisfaction is important in ascertaining information and providing a continuum of mental health services for all children.

As mentioned previously, demographic variables and other client characteristics have not been found to predict parent satisfaction. Rather, various studies have found that it is how parents experience their child's treatment which is associated with parent satisfaction. In their review of parent satisfaction, Young et al. (1995) reported that parents of children with SED emphasized the importance of "professional interpersonal skills and a coherent system of care" and that parent satisfaction was significantly correlated with perceived parent collaboration. Parents who reported dissatisfaction indicated the need for better communication with parents and a greater degree of parental involvement (Young et al.).

In the development of the Family Satisfaction Survey (FSS), Measelle et al. (1998) worked with a parent focus group that revealed four major areas of parent concerns including professionalism, job-related competencies, commitment to partnership with parent, and respectful, non-blaming view of parents. The resulting two factor survey for case management services included Interpersonal qualities/partnership practices which accounted for 82% of the variance and Job Related Competencies which accounted for only 6.2 % of the variance. Measelle et al., found that increased contact with parents was significantly correlated with parent satisfaction, whereas length of service or caseload of the caseworkers was not related. In the work of Riley et al., (2005) on the Youth Services Survey for Families, the highest level of responses (90%) were from staff being respectful and speaking to parents in a way that they understood. Harrington Godley et al., (1998) found that satisfaction scores were most highly correlated with individual counseling than any other service modality (group, social skills, etc.).

During the past decade, the importance of parent satisfaction has led to increasingly specialized measures. Kapp and Vela (2004) created the Parent Satisfaction with Foster Care Services Scale (PSFCSS) to gather feedback from these often 'unheard' clients. The PSFCSS consists of 5 scales including a) private social worker competency, b) state social worker competency, c) cultural competence, d) empowerment/client rights and e) outcomes of agency quality. The researchers found that basic social work/interpersonal skills such as respect for client values, preparing the client for meetings, including the client in decisions, and respecting client culture were all more likely to be associated with satisfaction.

King et al. (2001) investigated the major elements of parents' satisfaction and dissatisfaction with pediatric rehabilitation services. One of the more important aspects of this study was the researchers compared relatively satisfied to relatively dissatisfied parents, rather than looking only at general satisfaction. Not surprisingly, the highly satisfied parents mentioned elements of process, particularly respectful and supportive care. Relatively dissatisfied parents mention both process elements, such as respect and continuity of care, and structural elements, such as lack of access to services. Both the highly satisfied and relatively dissatisfied parents made comments about process, suggesting that "elements of professional caregiving process are more important to parents than are structural elements" (p. 126). In particular, King et al. noted that even relatively dissatisfied parents were able to 'see the positives' and noted positive process, while it appeared structural issues, such as access, may be a trigger for dissatisfaction. "Studies have identified practitioners' interpersonal skills (respectful and supportive care) to be one of the most important determinants of client satisfaction" (p 127).

Parent satisfaction is also being investigated with early childhood intervention programs (Fantuzzo, Perry, & Childs, 2006). Most studies now appear to recognize the multidimensional nature of parent satisfaction and the role of parents' experiences with various aspects of a program. Summers et al. (2005) focused on family-centered practice and the partnerships between parents and staff in special education services. The use of the term 'partnerships' indicates a shift in the language used when referring to parent participation. Summers et al. found a two-factor model of parent satisfaction: Child Focused (alpha = .92) and Family Focused (alpha = .91) services. The Child-Focused factor included reliable and competent staff who treated the children with respect and had the child's best interests at heart. The Family Focused factor included respectful and supportive interactions between parents and staff, such as clear communication, equality, and respect for the family's values. Summers et al.'s conceptualization of partnership focuses on the interpersonal skills which have been found to influence general satisfaction such as respect, commitment, open communication, friendliness, communication skills, and interpersonal factors including sensitivity to parents and clarity of communication.

These studies point to the importance of the parent relationship with the clinician/service providers, the interpersonal skills of practitioners, and the need for respectful collaboration with parents. "Research findings have consistently shown that the most important factor contributing to satisfaction in the healthcare context has been interpersonal relationships between staff and consumers" (Gerkensmeyer & Austin, 2005, p. 62,). Shifting the focus of parent satisfaction research from 'services' to the actual service providers (staff, case managers, clinicians) "personalizes research and highlights

the individual professionals who are generally considered to be the most important elements in service provision" (Young et al., 1995, p. 227). By investigating the experience and support parents perceive when seeking mental health services for their children, researchers can begin to give service providers meaningful program feedback about respectfully collaborating with parents to help ensure better support for both the child and family struggling with child/adolescent mental illness.

Methodological Issues

In order to understand the development of the client/parent satisfaction literature, it is essential to understand the methodological challenges in this area. Most of the issues revolve around sampling, data collection, and the underlying conceptualization of the construct of satisfaction. Therefore, the issues for both the client and parent satisfaction areas are similar and are combined in this section.

Sampling and Data Collection

The largest issue in the satisfaction literature appears to be the high levels, or 'ceiling effect' of satisfaction reported by clients/parents (Attkisson & Zwick, 1982; Essex et al., 1981; Harrington Godley et al., 1998; Larsen et al., 1979; Lebow, 1982; Riley et al., 2005; Young et al., 1995). This issue with the validity of reported satisfaction has been explained by the halo effect, social desirability bias, and lack of variance due to sampling bias (Harrington Godley et al.; Larsen et al.; Lebow; Riley et al.; Young et al.). The social desirability bias may contribute to high satisfaction scores because "parents may be eager to appear grateful, and could be nervous about offending mental health professionals" (Young et al., p. 225). Other researchers have cited cognitive dissonance as a reason for consistently high satisfaction scores (Bodin et al., 2007; King et al., 2001).

They argue that parents who are highly invested in a service or program, especially one they are 'free' to choose, are therefore less likely to indicate that the program is somehow inadequate. Lebow pointed out the 'reactivity' associated with survey methods and that specific steps, such as having non-practitioners administer the surveys, should be used to lessen the reactive problem. However, even the high levels of reported satisfaction could be compared to a baseline, but such norms across settings and services have yet to be developed (Harrington Godley et al.; Larsen et al.; Lebow).

Although social desirability is a criticism or caution noted in almost every satisfaction study, only Gaston & Sabourin (1992) formally tested this hypothesis using the CSQ-8 and the Marlow Crowne Social Desirability Scale (MCSD). Similar to the results of other satisfaction studies, satisfaction was not significantly associated with age, education, income, gender, therapist gender, or treatment modality. General satisfaction was also not significantly related to the social desirability scores, even when hierarchical regression was used to test therapist gender, length of treatment, and modality as moderators between social desirability and general satisfaction. The two aspects of the study significantly associated with overall satisfaction were length of treatment (more weeks of treatment associated with higher satisfaction) and alliance with the therapist as measured by the California Psychotherapy Alliance Scale (CALPAS). Thus, the findings of Gaston and Sabourin provide direct evidence that the relationship with the therapist was the most significant predictor of overall satisfaction ($r = .65$). What has been assumed to be social desirability bias resulting in high satisfaction scores may actually be a result of the importance of the client/practitioner relationship.

The work on the Assessment Questionnaire – 2 by Finn, Schroeder, and Tonsager (2004) also provided evidence that the social desirability was not significantly correlated with the four factors of new understanding, positive mirroring, positive assessor relationship and negative feelings. Finn, Schroeder, and Tonsager also correlated the hierarchical Parent Experience factor with social desirability and found a nonsignificant relationship. Thus, although often cited as a problem with satisfaction scores, very few researchers have actually investigated the relationship between social desirability and satisfaction. Contrary to what is often assumed about the cause of the ceiling effect in satisfaction scores, it appears that neither general satisfaction nor client experiences with assessment are significantly influenced or related to social desirability.

The satisfaction literature has struggled with sampling bias and the fact that those least satisfied with services are more likely to terminate early and not respond to inquiries about satisfaction (Larsen et al., 1979; Lebow, 1982). More recently, researchers have been using client information files to compare basic demographics (race, sex, education, etc.) to check for significant differences between 'responder' and 'nonresponder' sample groups (Measelle et al., 1998; Riley et al., 2005). Although this comparison helps ensure that the self-selected sample groups are not significantly different, demographics are not good predictors of satisfaction. Thus, although the groups may be demographically similar, that does not mean they would be similar in terms of satisfaction. Research has found that 'mutual termination' is more highly correlated with satisfaction than length of treatment (Lebow) so it would be important for researchers to try and gather data at the termination of all clients (mutual or not) and gain insight into the different experiences of those who terminate early. Although the demographic check should become a standard

part in comparing survey results, it is only the first step in a more standard survey analysis procedure.

The effect of early termination is most apparent when data are collected via cross sectional time periods. Studies have varied the length of time from two weeks to months, but regardless, the sample then includes clients who have just started treatment, those in the middle, and those post treatment, and excludes those who have terminated early. Thus, the more satisfied clients are most likely over represented in cross sectional data (Byalin, 1993). Suggested remedies have included more costly time series and longitudinal collection methods (Larsen et al., 1979; Young et al., 1995). One of the main reasons for using cross sectional data collection is the lowered time and cost when paired with the typical format of a mailed questionnaire. Unfortunately, mailed questionnaire data have a low response rate ranging from 19% (Essex et al., 1981), 28% (Byalin), 33% (Young et al.), 37% (Riley et al., 2005) to 52% (Gerkensmeyer, Austin, & Miller, 2006), with an average of 46% (Lebow, 1982). The cross-sectional and mailed survey data collection methods reflect the limited resources most community health centers and researchers have to collect data. Phone interviews have been used, but mostly as a follow up or when seeking child or adolescent data, where reading ability is even more of a concern than for adults (Shapiro, Welker, & Jacobson, 1997; Young et al.). Besides the low time and cost of mailed questionnaires, the format ensures that actual service providers are not giving the surveys, which should help decrease social desirability bias. Although cross sectional data may provide more variance, in that those at the beginning of treatment may have lower satisfaction scores than those towards the middle or end, studies have not differentiated how ratings may progress from the beginning to end of

treatment. Do those who remain in treatment have high satisfaction scores throughout? And are there different experiences for those who terminate after one session versus those who may terminate midway through treatment? To understand the effects of early termination and sampling bias in cross-sectional data collection, researchers must focus on the experiences parents and other clients have with practitioners (collaboration, alliance, respect) that most likely affects satisfaction and hence, service utilization.

Fortunately, researchers have made progress in terms of generating larger samples across varied settings and services. Harrington Godley et al., (1998) modeled a study using standardized measures and procedures across 22 publicly funded agencies within a state measuring 12 services (case management, crisis intervention, social skills, etc) and then compared agency vs. region *z*-scores so that agencies could clearly see if they were above or below the average satisfaction response in eight different categories. Using Medicaid youth, Riley et al. (2005) were able to sample 14 different community health centers within a state with a total of 534 surveys returned, and Gerkensmeyer and Austin (2005) used 5 very different sites including a wrap-around community site, state operated inpatient program for children, a non-profit hospital, home based intervention, and a state inpatient hospital for boys.

The results for Gerkensmeyer and Austin (2005) differentiated among the settings, with the two community programs reporting higher levels of parent satisfaction, decision making, and informing parents. Gerkensmeyer and Austin also found that satisfaction was significantly lower for parents whose children were not living at home, which may reflect higher levels of severity and different experiences for parents working with inpatient programs versus community health centers. It is important that researchers

continue to look at satisfaction across different types of services and settings (Harrington Godley et al., 1998; Lebow, 1982; Young et al., 1995) as clients may have different experiences that could help improve programs at various levels of mental health services. The majority of research has taken place at Community Mental Health Centers, a potential problem because those at a public facility have "little choice of facility, type of treatment, or practitioner" (Lebow, p. 284), which may also help explain high levels of reported satisfaction: clients may not be aware of alternatives or hold low standards for treatment.

Theory of Satisfaction

One of the earliest and continuing problems with the satisfaction literature is the lack of theory surrounding the construct of 'satisfaction'. Lebow (1982) pointed out that "to some, satisfaction means a minimum state of acceptability of services, whereas for others it means near perfection" (p. 247). Pascoe (1984) accurately pointed out that most satisfaction research implicitly uses a discrepancy model, where satisfaction is the "difference between actual outcome and some other ideal outcome" (p. 186). One disadvantage of a discrepancy theory is that it assumes any deviation from what is expected, positive or negative, results in dissatisfaction.

To determine satisfaction based on a discrepancy theory, a client's 'expectations' of service need to be addressed and defined. As Bailey and Simeonsson (1988) point out "parents may report satisfaction with a program simply because they know of no better alternatives or are comparing the service to no service at all" (p. 10). Through a series of interviews with clients, Williams, Coyle, and Healy (1998) attempted to delineate how clients come to their 'evaluations' of mental health services. The researchers found that

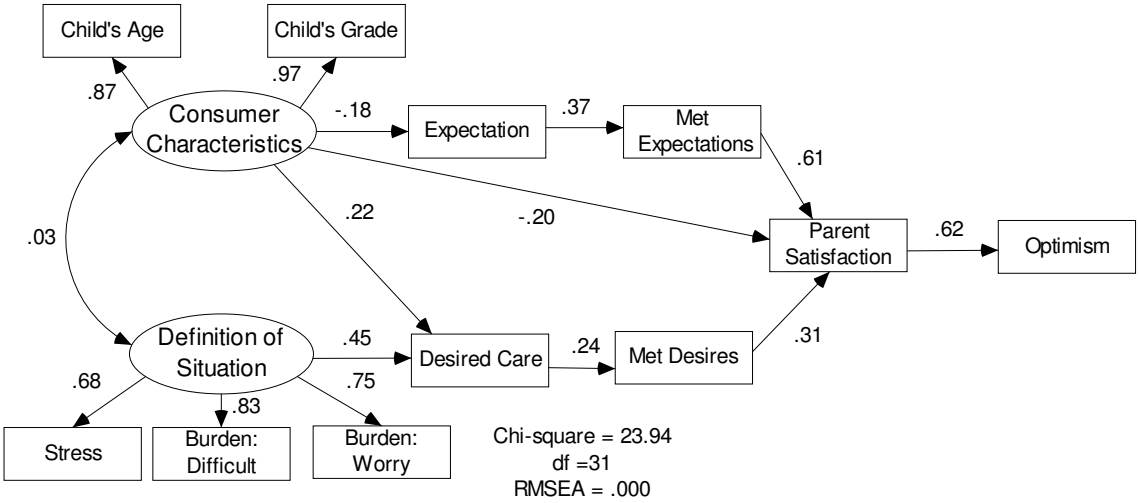
clients often had negative experiences related to services, but did not report dissatisfaction because they did not 'blame' the service provider. Duty, as defined by Williams et al., is how clients determine what a service should or should not do; it is somewhere between 'realistic' and 'idealistic' expectations. If the client does not believe the service provider was responsible for a negative experience, if it has not "failed in its duty", than the client may report being satisfied despite a negative experience.

The researchers also found that, even if clients thought the service provider was responsible for a negative experience, clients would often consider mitigating circumstances that would excuse the negative experience. Williams et al. (1998) labeled this Culpability and found that the most commonly excused failure of duty was the inability to 'cure' a client's mental health problem. Clients would either take personal responsibility for their illness or state that the therapist "was trying their best" and so it was not the therapist's fault the client was not 'cured'. Clearly, this work indicates that high levels of satisfaction do not mean there was an absence of negative experiences; they are not mutually exclusive. Rather, clients report satisfaction despite negative experiences when the client determines the service provider is not responsible. "Effort must be put into designing methods of accessing patients' experiences of service and the meaning and value they attach to them, whether these are positive or negative and whether they can be improved" (Williams et al., p. 1358). Therefore, in order to improve services and receive helpful feedback from clients, measures should move beyond general satisfaction and provide an avenue for measuring both positive and negative experiences.

In a review by Gerkensmeyer and Austin (2005) of the 34 parent satisfaction studies, none had a conceptual framework and only 2 presented conceptual definitions of parent satisfaction. Gerkensmeyer and Austin (2005)/Gerkensmeyer et al., (2006) used a discrepancy model in defining satisfaction, and attempted to measure pre-service expectations. Generally, studies only ask if expectations have been met, yet do not measure client or parent initial expectations at the beginning of services. Satisfaction was defined as "the difference between perceived services and consumers' desired and expected services" (Gerkensmeyer et al., p. 66). The researchers presented a model measuring consumer characteristics, consumer definition of situation, desired services, and expectations that when combined with actual service should influence level of met desires, met expectations, and thus, client satisfaction. Structural Equation Modeling (SEM) was used to test the hypotheses between the various observed and latent variables in the model as depicted in Figure 1.

The revised model indicated that neither consumer characteristics (child age/grade) nor parent definition of situation (stress, child difficulty, child worry) significantly predicted expectations. Rather, both consumer characteristics and definition of situation (DOS) contributed to desired care. Thus, parents related their situation to what services they desired, but it is unclear where parents gain their expectations. It may be parents have low expectations due to their lack of control when deciding on services or being unaware of other options (Lebow, 1982). However, as predicted, met expectations and met desires significantly predicted parent satisfaction, with factor loadings of .61 ($p < .01$) on met expectations and .31 ($p < .01$) for met desires (Gerkensmeyer et al., 2006).

Figure 1: Revised structural equation model for parent satisfaction. (Gerkensmeyer, et al., 2006)



The most interesting addition to the model was a parental optimism variable that was significantly predicted by parent satisfaction (Gerkensmeyer et al., 2006). The authors conceptualized that parent satisfaction can be both an ultimate outcome relating to service evaluation and an instrumental outcome influencing "engagement with the therapeutic regime, clinical outcomes, and quality of life outcomes" (Gerkensmeyer et al., p. 74). It may be that increased parent satisfaction, and hence increased optimism are two of the enabling variables related to further utilization of services/recommendations.

Developments on Dissatisfaction

Rather than continuing to confirm high levels of general satisfaction, some researchers are advocating the reduction of dissatisfaction as more informative. As Stallard (2001) points out, the positive scores consistently found with general satisfaction makes them an attractive option for service providers to obtain 'hard' evidence of good service quality. However, as previously mentioned, high satisfaction scores are not associated with service outcomes and the results of surveys, ostensibly to improve

practice and involve clients in shaping services, are rarely acted upon. Thus, Stallard suggests the active seeking of dissatisfaction to facilitate quality improvement of mental health practices.

Other researchers have begun to address dissatisfaction by dividing respondents into 'highly satisfied' and 'relatively dissatisfied' groups. Kapp and Vela (2004) considered those respondents one standard deviation above the average satisfaction score as 'highly satisfied' and all other respondents were considered relatively dissatisfied. King et al. (2001) argues that satisfaction and dissatisfaction are not one continuum, as the occurrence of one does not exclude the existence of the other. They considered 'highly satisfied' those respondents with a perfect score (32) on the CSQ-8 and those with a score of 23 or less as relatively dissatisfied. This indicates that dissatisfaction scores are often considered those with a mean less than 3 on a 5 point Likert scale. King et al. found that structural elements of service, such as access, appeared to be a particular trigger for dissatisfaction, while the interpersonal processes were associated with high levels of satisfaction.

The work of Williams et al. (1998) with client interviews demonstrates that high levels of satisfaction do not mean that clients do not have negative experiences. Stallard (2001) cautions that high rates of satisfaction do not indicate that the best services have been provided. Rather than continually lamenting the high scores constantly associated with general satisfaction, it appears researchers may be able to compare relative groups of satisfaction and dissatisfaction to help facilitate program feedback and improvements.

Importance of the Interpersonal Process and Collaboration

As noted in the previous sections on client and parent satisfaction, researchers are increasingly conceptualizing satisfaction as a multidimensional area most highly influenced by interpersonal experiences during services (Brannan, Sonnichsen & Heflinger, 1996; Essex, et al., 1981; Fantuzzo et al., 2006; Garland et al., 2007; Gaston & Sabourin, 1992; Gerkensmeyer & Austin, 2005; Gerkensmeyer et al., 2006; Greenfield & Attkisson, 2004; Kapp & Vela, 2004; King et al., 2001; Lewis, 1994; Measelle et al., 1998; Pascoe, 1984; Sheppard, 1993; Summers et al., 2005; Young et al., 1995). Previous focus on general client/parent satisfaction has made the feedback generated by research of little applicable use for program revision (Gerkensmeyer & Austin; Gerkensmeyer et al.; Young et al.). By shifting to the interpersonal aspect of child mental health services, researchers and evaluators can focus on the experiences parents have with child mental health services and improve specific areas, including friendliness of staff, warmth, and ways to collaborate and involve parents in decision making.

The growing importance of parental collaboration in the satisfaction literature is cited by Young et al., (1995) with the philosophy that "collaboration empowers parents and allows them to serve as more effective agents for assuring the quality of services their children receive. Satisfaction research is one dimension of the effort towards collaboration and signifies the willingness of concerned parents to advocate for their children and of providers to hear parental concerns" (p. 223). Fortunately, this collaborative spirit already exists in TA with children and TA can further anchor theory about what underlying mechanisms create positive experiences for parents that lead to greater parental engagement and optimism.

Psychometrics and Scale Development

Beyond difficulties with the sampling methods for satisfaction are the problems of reliability and construct validity of satisfaction measures. The typically high rate of 70-90% satisfaction is often attributed not only to sampling and response bias, but also to scale design rather than true perception (Kaufman & Phillips, 2000; Riley et al., 2005). In particular, the over reliance on Principal Components Analysis and lack of rigorous scale construction have been glaring issues in the satisfaction research. Thus, this section will review the primary psychometric issues with satisfaction scale construction and discuss more appropriate methods of scale development, specifically, confirmatory factor analysis and testing for invariance.

Item Development and Response Format

One shortcoming in satisfaction scale development includes not actively seeking the client perspective for satisfaction – thus the developed questionnaires may lack face and construct validity without client input (Measelle et al., 1998; Young et al., 1995). Researchers have suggested using focus groups with clients/parents to gain the perspective beyond that of the researcher's literature review (Measelle et al.; Young et al.). In the development of the Family Satisfaction Survey by Measelle et al. for case management, a parent support group gave four main areas (professionalism, competency, commitment to parent partnership, and respectful/non-blaming attitude towards parents) that indicate satisfaction is linked to how the parents feel treated by mental health professionals more than professional skill and competency. The inclusion of client perspectives may help researchers identify what actually contributes to variations in

satisfaction and positive experiences since demographics and outcome measures have been unsuccessful predictors.

Researchers must also be aware of the response format and how this may influence self-report responses. Most self-report surveys use a 4 or 5 point Likert scale with positively stated items (Gerkensmeyer & Austin, 2005; Measelle et al., 1998; Riley et al., 2005; Young et al., 1995). The current literature review found only one scale with a reverse scored item (Riley et al.) and a lack of items with different levels of difficulty that could check for extreme response patterns. It appears that in the interest of brevity, satisfaction questionnaires have consistently tried to reduce the number of items to the fewest possible, so that those reviewed were between 8 and 15 questions long with Riley et al. having the most items at 26 questions. However, the simplicity of the satisfaction questionnaires excludes components of better developed self-reports, such as extreme response detection (all positive, all negative) or social desirability bias subscales. Items describing different levels of satisfaction could result in increased variability among response with some representing a basic level of 'satisfaction' whereas others could correspond to exceeded expectations that could be informative about parent experiences.

Harrington Godley et al. (1998) note the importance of including a comment block as self-report measures may not cover every aspect of service that clients would like to comment. Many of the comments by parents mirrored other research findings that the parent and child relationship with the assessor and the support provided by the services (Harrington Godley et al.). Besides logistical concerns (more time slots, convenience to home) parents suggested more time with the clinicians and being treated equally/respectfully as areas needing improvement. Thus, although the Likert format

itself is widely used and supported, satisfaction surveys would do well to build in features that can prevent and detect response bias, as well as give clients the chance to comment on aspects not covered by the questionnaire. Also, instead of ignoring the negative feelings parents might have during their child's treatment (guilt, frustration, confusion) or that may be induced by how the parent experiences services (not listened to, disrespected) questionnaires should investigate these responses with the goal of reducing parent negativity experienced with child treatment.

Use of Principal Components Analysis

Almost all studies concerning client/parent satisfaction cite the lack of high level analysis for the scales developed or unknown psychometric properties – some do not even report reliability, and give only means and standard deviations (Attkisson & Zwick, 1982; Harrington Godley et al., 1998; Lebow, 1982; Riley et al., 2005; Young et al., 1995). Kaufman & Phillips (2000) found that sample sizes are often too small to meet the requirements for factor analysis, only 11% of satisfaction surveys tested inter-item reliability, and only 5% used factor analysis.

Even when the developers of previous satisfaction surveys attempted to use factor analytic techniques, the majority used Principal Components Analysis (PCA). Of the satisfaction measures directly reviewed, 10 reported using PCA and 2 reported using factor analysis, but did not report type (principal axis, maximum likelihood, common factor analysis) which makes it unclear whether PCA was used in place of factor analysis. Only four studies explicitly used Principal Axis, with one using Confirmatory Factor Analysis (CFA) in a second study. This disparity reflects a lack of understanding in the theoretical differences between PCA and Exploratory Factor Analysis (EFA) which needs

to be corrected. According to Fabrigar, Wegener, MacCallum, and Strahan (1999) "many researchers mistakenly believe that PCA is a type of EFA [Exploratory Factor Analysis] when in fact these procedures are different statistical methods designed to achieve different objectives" (p. 275).

Exploratory Factor Analysis (EFA) is designed to "uncover latent psychological attributes that account for correlations among observed variables" (Reynolds & Keith, 2009, p 20). Factors are latent variables which cannot be directly measured, but are the psychological constructs conceptualized to influence performance, attitudes, and behaviors. Questionnaires designed to measure client satisfaction assume that an underlying construct, or latent variable, of 'satisfaction' will influence how a client responds to questions concerning services. If client responses are consistent across items designed to measure 'satisfaction' than there should be high correlations among these responses representing the underlying construct.

Factor analysis is based on a common factor model, and measures both common and unique variance among variables. Common variance represents the underlying common factor that is consistent across measured variables, whereas unique variance includes both error and the unique aspects of a variable distinct from common variance. What separates PCA from EFA is that PCA does not differentiate between common and unique variance. Thus, rather than exploring latent variables, PCA is a data reduction technique that results in composite variables, or "weighted sums of the original variables" (DeVellis, 2003, p. 130). Although PCA maximizes the variance among measured variables, it does not explain the correlations among them (Fabrigar et al., 1999;

Reynolds & Keith, 2009). In other words "factors determine how items are measured, whereas components are defined by how items are measured" (DeVellis, p. 130).

Often, PCA and EFA will produce similar results, making it easy to misuse PCA when a researcher is really interested in underlying psychological constructs and EFA is more appropriate. In particular, if the unique variance for variables is low, than PCA and EFA will have congruent results (Fabrigar et al., 1999). However, the primary disadvantage of PCA is that by not accounting for unique variance it ignores the error which is present in correlation (covariance) matrices of variables (Kline, 1994).

Another area of difference between PCA and EFA is in the identified correlations among factors. Fabrigar et al. (1999) provide evidence that correlations between factors using an oblique rotation of PCA were much lower than using the same rotation in a factor analysis. Fabrigar et al. explain that "it makes sense that PCAs should generally underestimate relations among the constructs, because random error is included in the components. Because factor analyses remove random error from the factors, the relations among factors are more likely to approach the population values" (p. 289).

Researchers using PCA or EFA have a choice between using orthogonal rotation, where factors/components are uncorrelated, or oblique rotations, which allow correlations between factor/components. Of the satisfaction scales reporting type of rotation used, half reported using orthogonal techniques and half reported using oblique rotations. Researchers may choose orthogonal rotations (often Varimax) in order to create 'clean' results and simple structure, yet may not realize that oblique rotations do not require factor correlations, and can produce uncorrelated factors if that best represents the data. As Fabrigar et al. (1999) point out, that for many, if not most, constructs in psychology

such as traits, attitudes, and cognitive skills, there is theoretical and empirical evidence that constructs will be correlated. Thus, restricting correlation among factors may lead to less accurate or unrealistic results, while understanding correlations among factors may promote better conceptual understanding.

In general, "PCA should not be used as a substitute for EFA" (Fabrigar et al., 1999, p. 283) when researchers are interested in underlying constructs, such as 'general satisfaction'. Although there appears to be a trend in the development of satisfaction scales towards using EFA, with more recent studies moving away from PCA, it appears there is still quite uneven distribution of knowledge about the application of scale development using factor analysis. The earliest EFA study reviewed occurred in 1981 (Essex et al., 1981), yet PCA has been used as recently as 2004 (Kapp & Vela, 2004). For researchers creating a new scale without a strong underlying theory, EFA is an appropriate technique, but one that may require further review of basic standards.

Confirmatory Factor Analysis

In conducting confirmatory factor analysis (CFA) a researcher uses theory to specify the number of factors and patterns of factor loadings. The advantages of CFA is that is allows a researcher to test specific hypotheses about the data (Keith, 2006). Various fits statistics are used to evaluate the hypotheses and provide feedback about the model's ability to 'fit' with the data (Keith). Because of the a priori nature of CFA, researchers are less likely to capitalize on chance findings (Fabrigar et al., 1999). "CFA is commonly used in psychological assessment research to address questions related to the measurement of psychological constructs and construct validity" (Reynolds & Keith, 2009). Although CFA is not commonly used in scale development for client satisfaction

research, it provides a useful method to test hypotheses and refine measures based on theory, rather than using the more data driven approach of EFA.

Given the reliance of CFA on specifying the relationships among factors (latent variables), measured variables, and factor loadings, the importance of sound theory cannot be overstated. The greater freedom associated with CFA, also allows for "more opportunities for making poor decisions" (DeVellis, 2003). Just as EFA processes can be misused, intentionally or not, so can CFA. "No matter how sophisticated our analyses, they cannot turn bad data into good or a poor design into a powerful one" (Keith, 2006, p. 381). Just because one model 'fits' the data does not mean another model might not fit as well or better. Since all models cannot be tested, the importance of a sound rationale is essential to the appropriate use of CFA (Kline, 1994).

The evaluation of models in CFA is based on a series of fit indices designed to represent the 'fit' of the data to the specified model. The most commonly used fit statistic is χ^2, which compares the actual and implied covariance matrices. A nonsignificant difference between the two matrices implies that the model is consistent with the data and may "approximate reality" (Keith, 2006). The degrees of freedom are used as an indicator of parsimony in a model, in that the more degrees of freedom, the fewer paths have been drawn (or more values are constrained to zero). In order to compare competing models, a researcher can use χ^2 to test for a significantly worse fit between nested models. Nested models are models that can be obtained by deleting paths, or constraining loadings to zero, making a more parsimonious model. Although researchers value parsimony, the increase in df (deleting paths to create a nested model) is not 'worth' a significantly worse fit between the data and the implied model. In the

case where the difference in χ^2 does not indicate significantly worse fit, the more parsimonious model (with greater degrees of freedom) may be chosen.

One of the disadvantages of χ^2 is its sensitivity to large sample sizes. Adequate models may be rejected due significant differences in the implied and actual matrices which is a function of the sample size rather than lack of fit in the models. Other fit statistics commonly used and described by Keith (2006) include: the Goodness of Fit (GFI), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). CFA and other structural equation modeling techniques are generally considered large sample (500+) techniques. Older rules of thumb for factor analysis in general held that a ratio of 5 or 10 subjects per variable or item in scale development was necessary. However, Guadagnoli and Velicer (1988) and MacCallum, Widaman, Zhang, and Hong (1999) noted that these suggestions were not supported empirically. The stability of factors is also related to the ratio of the number of variables to the number of factors, with more variables per factor resulting in greater stability. Thus, the more overdetermined (at least 3 to 4 variables per factor) a factor, the better. In addition, as the communalities of variables, or their loadings increased (.60 and above), the influence of sample size decreased. MacCallum et al. conclude that the common rules associated with factor analysis are not invariant across studies, and rather the level of communality and overdetermination of factors have more influence than strict sample size on the recovery of population factors.

However, it is still necessary to have more subjects than measured variables, and in light of the previous findings, sample sizes of at least 100 may be adequate for CFA. MacCullum, Browne, and Sugawara (1996) demonstrate how to calculate power in

structural equation modeling for RMSEA utilizing sample size and degrees of freedom. In general, increases in sample size and degrees of freedom are associated with more power. The smaller the sample size, the more important the number of indicators per factor becomes; three or more is recommended (Keith, 2006).

The use of CFA in the development of satisfaction measures has been rare, with only one study using CFA as a follow-up to an initial EFA. Brannan et al, (1996) provides the only example of a comprehensive use of CFA in assessing the validity and reliability of the Satisfaction Scales. The Satisfaction Scales were designed to measure content areas, including a) access and convenience, b) child's treatment, c) parent services, d) family services, e) relationship with therapist, f) staff responsiveness, and g) global satisfaction (which consisted of 5 items from the CSQ-8). The sample consisted of 544 parent responses to their children's outpatient mental health services. The 29 items were grouped into 10 indicators used throughout the analysis, with each indicator the mean of at least two items. Each factor had a minimum of two indicators throughout the study; however, it is unclear why the individual items were not allowed to load directly onto the factors. The use of indicators reduced the factor–to-variable ratio and the indicators were less desirable composite scores.

The initial model consisted of 5 factors, each with two indicators, and the factors were allowed to correlate. The 5 factors were access and convenience, child's treatment process, parent and family services, relationship with therapist, and general satisfaction. Results for the initial model were factor loadings between .82 and .98 for all indicators, and correlations among factors ranging from .53 to .92. Due to the high factor correlation between child's treatment process and parent's relationship with therapist (r = .97) these

two factors were combined in a four-factor model. The resultant χ^2 of the four factor model was 82.50 (df = 29, p<.001). The difference in χ^2 between the two models was not statistically significant, indicating that both models 'fit' the data equally well. Therefore, the more parsimonious four-factor model was chosen as the final model. Brannan et al. (1996) also tested a three-factor and single-factor models, but these resulted in worse overall fit and were rejected.

Alternative models that were not tested by Brannan et al. (1996) included a hierarchical model with general satisfaction influencing the other four factors, or a nested model controlling for general satisfaction across the initial four factors. The Access and Communication factor dealt strictly with structural issues (convenience, promptness) related to services and the General Satisfaction factor consisted of the 5 questions from the CSQ-8. The Parent and Family Services factor asked only about the amount of and satisfaction with parent/family services. Noticeably, the Child Treatment/Relationship with Therapist factor was the only one which included interpersonal or process variables. Items from this factor could be conceptualized as collaboration (parent involvement, agreement, and communication with therapist during child's treatment) and parent-therapist relationship (responsiveness, ease of contact, therapist's communication about the child's condition). Considering the previous findings on client satisfaction, it is not surprising that in the initial five-factor model, the correlations between child's treatment and global satisfaction was .92, between parent relationship with therapist and global satisfaction was .90, and between the child treatment and relationship with therapist was .97. Other studies have repeatedly shown that the positive interactions and interpersonal

experiences are more closely related to general satisfaction than access or structural elements.

Brannan et al. (1996) provides an example of how the underlying factor structure of a parent satisfaction questionnaire can be evaluated and refined using CFA. In addition to testing specific hypotheses and comparing competing models of latent variables, CFA also allows researchers to compare the consistency and meaning of items across groups through measurement invariance.

Measurement Invariance

Measurement invariance is the "extent to which items or subtests have equal meaning across groups" (French & Finch, 2006, p. 379). Or, put another, way, those who have the same values of a latent trait should receive the same observed score, regardless of group membership (Reynolds & Keith, 2009). Determining measurement invariance is extremely important in test development, such as cognitive abilities or achievement tests, which have high stakes for test takers. Testing for invariance allows researchers to test for systematic bias. According to French and Finch a respondent's score should not depend on variance due to things other than the construct being measured, such as group (gender, ethnicity) membership.

In order to test for different levels of invariance across groups, multiple group mean and covariance structure analysis (MG-MACS) is used to compare different aspects of measurement across multiple groups using confirmatory factor analysis. The first step is to determine the baseline factor structure, or configural invariance, for two groups. In short, the same factor model should be established for each group, with the same number

and pattern of factor loadings. “This model is analogous to running independent factor models…and then combining the information” (Reynolds & Keith, 2009).

After configural invariance has been achieved, with good model ‘fit’, other parameters are increasingly constrained to be equal across groups. After each additional level of constraint, the difference in χ^2 and degrees of freedom (along with other fit statistics) may be used to determine if there is significantly worse ‘fit’ in the more constrained model.

Factor loading or metric invariance is crucial, though not sufficient in and of itself, to determine measurement invariance. To test for metric invariance, the unstandardized factor loadings are constrained to be equal across groups. If a statistically significantly worse fit for factor loading invariance is found, than the items or subtests “are not measuring the factors in the same way” for the different groups (Reynolds & Keith, 2009, p. 42).

The next level of constraint is intercept invariance, where the measurement intercepts are set equal across the groups. According to Bryne and Stewart (2006) intercept invariance allows for comparison of latent factor means. Or, researchers are testing “whether differences in the factor means can account for the difference in the observed mean of a subtest” (Reynolds & Keith, 2009). In other words, without this level of invariance, one cannot assume that differences in observed means reflect actual differences in latent means. Although it is often assumed that observed scores reflect true differences in the latent construct of interest in (i.e. satisfaction scores reflect true differences in overall satisfaction) this assumption is rarely tested empirically for attitudinal scales.

Residual invariance involves constraining the measurement residuals (or errors) to be equal across groups. If residual invariance holds, this constitutes 'strict' factorial invariance, which implies that any group differences in measured scores are solely attributable to true group differences in the latent variables (Bryne & Stewart, 2006). However, even if residual invariance is not found, factor means can still be compared, assuming both factor loading and intercept invariance has been found. When strict factorial invariance is found, then researchers can be assured that scores from a test are unbiased between the two groups compared; scores do not depend on group affiliation (Reynolds & Keith).

Once measurement invariance has been established between groups, than differences in the latent means, variances, and covariances can be tested for the groups. For example, in cognitive abilities testing, once it is determined items are not biased towards girls or boys, then comparisons in latent abilities are possible. By comparing latent means, Reynolds, Keith, Patel and Ridley (2008) found slight advantages for females in long-term retrieval, and short-term memory, while males showed advantages in visual-spatial reasoning and verbal comprehension. It is important to note that the use of latent means offers a significant advantage in determining 'true' differences in abilities or latent variables, as once again, the use of composite scores could provide different results, as they do not account for error and unique variance (See Reynolds et al., 2008).

Without empirical evidence of measurement invariance, researchers relating differences in scores from a measure may incorrectly assume that the measure is unbiased and functions similarly across groups. Testing for measurement invariance is an important step in assuring that a test or questionnaire functions similarly for different

groups of respondents and can be used to test for commonly found groups, such as age, sex, ethnicity, and education level.

Although most often used in test development, measurement invariance can also be used to establish congruence between different translations of an instrument, such as Bryne and Stewart's (2006) instructional use of MG-MACS to compare Chinese and English versions of the Beck Depression Inventory II (BDI-II). In this study, configural invariance was established across both groups. Factorial invariance was found for all items except one (fatigue) while intercept invariance did not hold for six items (constraints) across the two groups. The single item without factorial invariance is of the greatest concern, as it implies the 'fatigue' item is not functioning similarly between the Chinese and American adolescents. Bryne and Stewart interpret the lack of factorial invariance for the fatigue item to indicate that the item was more ambiguous for the Chinese respondents, who had lower scores on that item, than the American respondents. This may imply "important differences in the theoretical structure of the underlying construct of the Somatic Elements facet of adolescent depression" (Bryne & Stewart, p. 312). Thus, testing for measurement invariance is not limited to test development and can provide important feedback about item functioning across groups in questionnaires.

In general, scale development in the client and parent satisfaction literature suffers from methodological problems concerning sample bias, sample size, and lack of explicit theory. Most importantly, it lacks psychometric rigor and has inconsistent use of appropriate techniques, most notably a default use of PCA rather than EFA. With the rare exception, the client satisfaction literature has not used more powerful techniques, such as CFA and MG-MACS, to test theoretically driven models and hypotheses about the

factor structure and invariance of satisfaction measures across groups. Applying these techniques may facilitate greater understanding and provide more empirical evidence for scale development and investigation of constructs, such as interpersonal processes, that appear to be highly related to general satisfaction.

Summary and Statement of the Problem

Therapeutic Assessment is designed to be a positive intervention for adult clients and children, adolescents, and their families. The principles of TA encourage high levels of parent involvement, questioning, and interpretation, which can provide a positive change experience and a renewal of hope for parents thus leading to greater follow through on recommendations and parenting strategies for the child. The literature presented in this review indicates that TA-C can provide the theoretical structure and high levels of collaboration and parent involvement currently highlighted in the parent satisfaction research. Recent research is recognizing the important role that parents play in seeking and maintaining services for their child, and that parent satisfaction may influence how parents follow through on recommendations and treatment plans. It has also been found that parent satisfaction may be more related to interpersonal relationships with staff (collaboration, respect, feeling heard) rather than child outcomes; parents may value the idea of good 'care' instead of 'cure'. However, general satisfaction does not tell researchers and program evaluators what interpersonal experiences parents are having and what experiences (parent-assessor relationship, collaboration, systemic awareness, etc.) are most salient for overall parent satisfaction.

Studies from all areas of client, patient, and parent satiafaction have demontrated the importance of interpersonal processes in quality health and mental health delivery.

"Thus, a more collaborative approach to client-provider interaction, influenced by a common desire to make the experience more 'human, respectful, and understandable' to consumers (Finn, 2007, p. 5) has emerged in both the medical and pscyhological fields" (Tharinger, Krumholz, et al., p. 4). As TA-C continues to demonstrate, in order to help children, the clinician must also assist the parents by providing the opportunity for new conceptualizations, attributions, and experiences. By expanding beyond general parent satisfaction to parent experiences, researchers may be able to identify which interpersonal processes are most valued by parents and are essential for child mental health service providers to focus on in program delivery and evaluation. By creating a parent measure that explores the interpersonal experiences of parents during a child assessment, researchers will be able to simultaneously provide more verification of the collaborative techniques of TA, and investigate the experiences most related to parent satisfaction.

The purpose of this study is to design a parent questionnaire that measures the different experiences parents may have with an assessment, based on the underlying processes of the Therapeutic Assessment model. Some of the most significant short comings in the parent satisfaction literature have been related to methodological issues and the lack of appropriate scale construction techniques. Confirmatory factor analysis and multiple group mean and covariance structure analysis (MG-MACS) will be used to investigate the structure of a six-factor model of parent processes, test competing models, and provide a revision of the pilot version of the parent questionnaire. Thus, this study will attempt to provide clarity about the aspects and processes of TA-C which makes it a positive and transformative experience for parents, and provide direct evidence relating these parental experiences to general satisfaction.

Chapter III: Method

Development of the Instrument

The Parent Experience of Assessment Scale (PEAS) was developed to provide a more quantitative way to measure the underlying mechanisms in TA-C and investigate how salient these aspects are to parents during the psychological assessment of their children. In particular, the interpersonal process categories (parent-assessor relationship, and collaboration) discussed previously were hypothesized to be a major areas of which parent feedback and responses are relevant in expanding the interaction of these processes in TA. In addition, a child-assessor relationship category was included to measure parents' perceptions of how well the assessor worked with and engaged the child. Parents were also asked about negative interpersonal experiences during the assessment, as Williams et al. (1998) demonstrated that high satisfaction does not necessarily imply an absence of negative experiences. Thus, the PEAS hypothesized multiple factors and experiences of parents during a child assessment that warrant measurement and can provide useful feedback. The construction of the PEAS as well as the research questions and analysis for the current study are discussed below.

Item Generation

The Assessment Questionnaire-2 developed by Finn, Schroeder, & Tonsager, (2004) was based on adult client experiences with therapeutic assessment and provides a basic framework for conceptualizing different facets of the PEAS. The general categories of new awareness, positive relationship with the assessor, and negative feelings were carried over into the PEAS. Initial items and additional categories for the PEAS were generated by reviewing transcribed parent interviews following the

completion of a child TA through the Therapeutic Assessment Project (TAP). The interviews asked parents about the TA experience, and about their attributions for their child's problems, role of family, future outlook, skills, and information gained from the assessment. Potential items were also generated by research team members who had worked with and observed parents and children during TA. Initial categories included information about the child, new skills, new understanding of child, systemic views, feeling understood, child relationship with assessor, collaboration, parent relationship with assessor, negative feelings about the assessment, positive feelings about the assessment, and optimism/pessimism about the future.

The items within each category were constructed with both positive and negative wordings (introducing the need for reverse scoring). Negatively worded items are often used to counteract acquiescence or agreement bias (DeVellis, 2003). By including negatively worded items, parents will be unable to simply choose the highest or lowest score and thus should help create a more valid scale. Some items were also designed to be more difficult than others. For example, the item "I learned a tremendous amount about my child from the assessment" is more difficult to agree with than "I learned a lot from the assessment". The first item differentiates between those parents who found the assessment extremely insightful versus those who found it merely helpful or routine. According to DeVellis, "in general, very mild statements may elicit too much agreement when used in Likert scales" (p. 79). Therefore, making items more difficult to agree or disagree with may reduce the 'ceiling effect' so often associated with satisfaction measures and may reflect more variance in parent experiences. Items were also designed at a fourth-grade reading level and with natural language wording that avoids jargon.

The initial 78 items were then given to 9 expert judges to sort into non-overlapping groups based on item similarity. The judges were asked to provide names for the categories they determined and were not limited in the number of categories to create. This procedure was modeled by Tellegen (1981), and the results were entered into co-occurrence matrix, or "a symmetric similarity matrix in which each pair of items was assigned a similarity value equaling the number of judges who grouped the two items together" (p. 220). The co-occurrence matrix was then submitted to an exploratory factor analysis to determine which items were grouped consistently enough to distinguish preliminary factors.

Preliminary Factor Analysis

The results were analyzed using Principal Axis factor analysis with Direct Oblimin rotation (KMO = .929; Bartlett = 4523.450, df = 2906, p <.000). Solutions with 3 through 8 factors were analyzed using a .55 cutoff for minimum loading on a factor. Most factor loadings ranged from .63 to .81. Items that did not load on a factor or were highly loaded on multiple factors were removed. Some items that loaded on two factors were revised to measure only one factor, and if a factor did not have enough items, additional items were created. The preliminary measure consists of 64 statements divided among 6 subscales. The pilot subscales are described below:

- Parent-Assessor Relationship is designed to measure the interpersonal relationship between the assessor and parent, including feeling respected, valued, and heard by the assessor.
- Collaboration assesses how informed and involved the parent was during their child's assessment.

- Child-Assessor Relationship investigates the parent's perception of how well their child worked with and responded to the assessor.
- New Understanding of Child subscale highlights new information and awareness gained by parents about their child from the assessment.
- The Systemic Awareness subscale assesses how much the parent considers the family's role in helping/maintaining the child's problems. TA is designed to encourage parents to become a positive force in helping their child and re-energize parents who have 'tried everything'. The Systemic Awareness subscale is designed to probe the more systemic/family aspect of children struggling with mental health problems.
- Negative Feelings about the assessment include anxiety, guilt, and frustration that parents may feel, regardless of how collaborative the assessment is conducted. More responses to this scale may be related to increased severity of their child's symptoms, but especially high levels may reflect the 'blame' parents may often feel for their child's problems.

The six subscales resulting from the preliminary factor analysis were generally consistent with the original content areas for items. The new understanding of child subscale appeared to incorporate items from the content areas of information about the child, new skills, and new understanding of child. The content area of parents feeling understood was incorporated within the parent-assessor relationship items. Positive feelings about the assessment did not retain its own content area, whereas the negative feelings towards the assessment remained a distinct area. Lastly, the optimism/pessimism about the future items tended to load on multiple factors, as the

positive/negatively worded aspects of the items may have confounded their loading together as a general category about future outlook. Some of the items were revised to load more clearly on either new parenting skills (in new understanding of the child) or systemic awareness areas, while others were dropped due to low factor loadings. Overall, the major areas of child-assessor relationship, collaboration, parent-assessor relationship, systemic awareness, new understanding of the child, and negative feelings were retained across the initial analysis and form the six subscales of the pilot version of the PEAS

Implementation

Pilot testing of the PEAS with a community sample of parents/guardians who received a psychological assessment for their child was essential to further refine the subscales. In order to see how the experiences of parents vary across assessment type and setting, data from four different sites were included in this study. The sites included a local private assessment practice, a local neuropsychological assessment practice, a community child clinic, and a public school. This diversity of sampling allowed for responses from a variety of child assessments, ranging from traditional community and private practice to collaborative to TA research practice. Further descriptions of each site will be provided in the results section.

Participants

Participants in this study consisted of parents or legal guardians of children and adolescents who have received a psychological or neuropsychological assessment. Confidentiality was maintained by assigning each participant a numerical ID once the information sheet and questionnaires had been completed and given to the principal investigator for data entry. Spanish translations of the questionnaires and consent form

were provided to allow more diversity in parent/guardian participation. Ideally, each site would receive between 30 and 40 completed questionnaires. Due to the importance of sample size in the planned factor analyses, the minimum goal for the number of surveys was 125.

Instrumentation

The Client Satisfaction Questionnaire (CSQ-8; Appendix N): The CSQ-8 (Attkisson & Zwick, 1982; Larsen, et al., 1979) is the most widely used measure for general client satisfaction. Although originally normed for adult clients, it has more recently been used in parent satisfaction studies (Byalin, 1993; Gerkensmeyer & Austin, 2005). The CSQ-8 is single factor scale with high (.93-.96) reported reliability (Attkisson & Zwick; Gerkensmeyer & Austin). See Appendix N.

Parent Experience of Assessment Scale – Pilot Version (PEAS; Appendix O): The pilot version (64 item) PEAS was administered in all settings at the conclusion of the child's assessment. The preliminary scale consists of 6 subscales: Collaboration, Parent-Assessor Relationship, Child-Assessor Relationship, New Understanding of Child, Systemic Awareness, and Negative Feelings. The measure is based on a 5-point Likert scoring system, with some reverse scored items. The average rating for each subscale is calculated via an electronic scoring sheet. See appendix O.

Procedure

Each site was provided packets of the instruments and consent forms to be distributed to study participants. Parents/guardians were asked to complete the questionnaire and the 8-item CSQ as a check out procedure after the last assessment or

feedback meeting. The data collection occurred at the conclusion of services so parents were not asked to rate a clinician from whom they were still receiving services.

Because assessments are often shorter than other psychological services, it was expected that the problem of early termination so often encountered in other treatment services would not be as salient an issue. Rather, multiple options for enabling a good response rate were in place as parents/guardians could complete the measures electronically, on paper, or over the phone as site resources allowed.

Research Questions and Hypotheses

Research question 1

What is the factor structure underlying the PEAS?

Hypothesis 1. The PEAS will demonstrate a six-factor model, with each factor corresponding to one of the PEAS subscales.

Rationale. As cited earlier, the majority of client and parent satisfaction questionnaires have been developed using principal components analysis (PCA) which theoretically and statistically provides composite variables, rather than information about underlying factors (or latent variables). Confirmatory factor analysis (CFA) was used to evaluate the expected six factor model, based on areas and processes highlighted in Therapeutic Assessment. The use of CFA was appropriate given the a priori hypothesis of the PEAS pattern of factor loadings. However, competing models were investigated to test the adequacy and fit of the 6-factor model.

Research Question 2

Does the PEAS measure the same constructs across groups?

Hypothesis 2. It was expected that the items and latent variables of the PEAS would meet criteria for strong invariance across respondent type, language, and social economic status.

Rationale. In order to assume that differences in observed means reflect differences in latent means, rather than bias associated with group membership, strong (metric and intercept) invariance should be demonstrated. Although some items and factors may not be as applicable for all participants, and there may be true latent mean differences between groups, it was expected that the items as measurement tools would function similarly across groups. Respondent type (parent vs. guardian) group invariance was investigated to determine whether items functioned similarly across groups, or whether different versions of the PEAS were indicated. A Spanish version of the PEAS was utilized and testing for measurement invariance is essential in determining whether the translated items were functioning similarly to the original items (See Bryne & Stewart, 2006 for an example). The Spanish version was translated by a native Spanish speaker in the mental health profession, and was then backtranslated (from Spanish to English) by a second native Spanish speaking mental health professional to ensure consistency in interpretation. Lastly, research has not generally found differences in general satisfaction to be associated with demographics; however, one study found that education showed an inverse relationship with general satisfaction (Bodin et al., 2007). Although efforts have been made in the development of the PEAS items to ensure clarity and ease of reading, testing for invariance across SES would help ensure that different groups are interpreting and responding to the items in a similar fashion.

Research Question 3

How do various aspects of interpersonal relationships with a clinician, as measured by the PEAS, relate to overall parent satisfaction?

Hypothesis 3a. General parent satisfaction should be most highly related to the parent-assessor relationship, level of collaboration, and the parent's perception of the child-assessor relationship.

Rationale. Previous research in the parent satisfaction literature has already shown that overall parent satisfaction was more related to interpersonal experiences, including higher levels of collaboration with parents and parents feeling respected and heard by professionals (Gerkensmeyer & Austin, 2005; Measelle et al., 1998; Riley et al., 2005; Young et al, 1995). When reporting areas needed for improvement, parents cite better communication and a greater degree of parental involvement (Young et al.).

'Collaboration' as conceptualized and practiced in Therapeutic Assessment (and hence the PEAS subscale) includes the parents helping set the scope of the assessment, being informed about each step in the assessment process, contributing ideas about the validity of the test results, and working as a team with the assessor to help their child. Items similar to the Collaboration subscale can be found in the Participation in Treatment factor of the Youth Services Survey for Families (YSSF; Riley et al., 2005) and individual items (kept me informed, find the right services, included me in decision making) on the Parent Satisfaction Scale (PSS; Gerkensmeyer & Austin, 2005).

'Parent-Assessor Relationship' as designed by the PEAS includes feeling respected, liked, and listened to by the assessor, as well as a reciprocal relationship of the parent feeling close to the assessor, liking the assessor, trusting the assessor, and feeling

the assessor was genuinely interested in helping. These items are similar to the Dignified Treatment factor from the Client Satisfaction Survey (CSS) by Essex et al. (1981), items from the PSS (treated with respect, listened to what I had to say, support), and the Cultural Sensitivity factor of the YSSF, which also included respect for religious/spiritual beliefs and cultural/ethnic background (Riley et al., 2005). However, the Parent-Assessor Relationship items on the PEAS includes reverse scored items and, rather than only rating how the assessor treated the parents, it also allows the parents to rate how they felt about the assessor, thus providing information about the reciprocal nature of the relationship.

The 'Child-Assessor' subscale on the PEAS asks the parents how comfortable the child felt with the assessor, how well the assessor worked with the child, if the assessor and child both appeared to like each other, and if the assessor seemed to understand the child. Most often in parent satisfaction surveys, questions regarding the child focus on treatment outcomes (getting along better with others, better at school, daily coping, helped parent deal with child's problems, symptom reduction). However, parents still reported being satisfied, despite nonsignificant outcome findings (Plante et al., 1998). Therefore, it seems that the support that parents and children receive from treatment services that should also be measured. On the Youth Client Satisfaction Questionnaire (YCSQ) two factors emerged for the child – Relationship with Therapist and Benefits of Therapy, which were highly interrelated (interfactor correlation .61, $p<.0001$; Shapiro et al., 1997). Parallel to the parent satisfaction data, child reported relationship with therapist was significantly higher than child reported benefits of therapy (Shapiro et al.). Thus, parents are aware of their own relationship with the assessor, and they also have

important perceptions about their child's relationship with the assessor which need to be measured.

Based on the principles of TA and the importance of interpersonal relationships for parent satisfaction, it was hypothesized that these three interpersonal aspects would be the most highly associated with overall parent satisfaction.

Hypothesis 3b. Lower feelings of negativity associated with the assessment should be correlated with higher general satisfaction.

Rationale. The parent satisfaction literature has thus far phrased questions on parent surveys in a positive and neutral frame (Were you satisfied with….) and only qualitative comments allowed clients to express negative feelings or suggestions (Essex et al., 1981). As Gerkensmeyer & Austin (2005) pointed out, by assuming any negativity equals dissatisfaction, researchers have shown a lack of clear theoretical conceptualization. It is possible parents could be satisfied without significant change in outcome variables and may have suggestions for program improvement (Gerkensmeyer & Austin). Williams et al. (1998) demonstrated that high reported levels of satisfaction do not mean that clients did not have negative experiences with services. Rather, clients may not hold the service provider responsible, citing functional limitations of the therapist or system.

Rather than ignore negative affect that may accompany a child's assessment, the Negative Feelings about the assessment subscale asks parents about feelings of guilt, lack of parenting efficacy, feeling blamed, ashamed, or overwhelmed. As mentioned previously, the satisfaction literature has tended to ignore the fact that parents often feel blamed by service providers (Measelle et al., 1998; Young et al, 1995). By including this

subscale, the PEAS allows practitioners to gain an understanding of 'normal' or baseline level of negative affect generally associated with needing an assessment for one's child. This subscale also provides valuable feedback for program revision so that parents who feel blamed or anxious can receive more support in the future. Although some negative affect is expected regardless of how well the assessment process unfolds, it was hypothesized that lower levels of negative feelings toward the assessment should be associated with higher overall parent satisfaction.

Research question 4

Can the PEAS subscales differentiate between collaborative and traditional assessment practices?

Hypothesis 4. It was expected that general parent satisfaction and the PEAS subscale scores would be significantly higher for therapeutic assessments than traditional child assessments. Specifically, the Learned New Things and Systemic Awareness subscales are most likely specific aspects of a Therapeutic Assessment that should contribute to Therapeutic Assessment receiving the highest overall parent satisfaction scores.

Rationale. Although the goals of traditional assessment include gathering data that help to describe a client's situation accurately and informing treatment, the goals of TA go beyond making a diagnosis or explaining standardized scores to parents (Finn & Tonsager, 1997).

Collaboration is a central component of TA, and it was expected that participant responses will be significantly higher for collaborative assessment than traditional assessments. Research of parent satisfaction has noted the importance of including

parents in the treatment of their child. Summers et al.'s (2005) scale of family-professional partnerships demonstrates the collaborative nature between family members and professionals in a long term relationship, such as a child receiving special education services. TA strives to create a collaborative relationship by infusing assessment with aspects of therapy, such as 'cognitive empiricism' in order to help parents be active partners in the assessment. It was hypothesized that the greater level of collaboration in TA will be reflected in parent responses on the PEAS.

The New Understanding of Child subscale on the PEAS questions parents about learning new ways of interacting and responding to their child, changing their perception of their child, understanding child strengths, as well as gaining new information from assessment results. This scale speaks to the intervention nature of child TA in that the parents are able to question, assimilate new information, and receive tailored feedback in a supportive environment enabled by the interpersonal relationships in Hypothesis 1. Also, instead of focusing on child outcomes, the scale more closely investigates feelings of better parenting skills and effectiveness, new ideas, and new understanding. Although hypothesized to be related to parental satisfaction, the New Understanding of Child subscale is most likely not as highly correlated as parent-assessor relationship, collaboration, or child-assessor relationship to general parent satisfaction. Rather, scores on this subscale should increase with more collaborative types of assessment, such as TA versus traditional assessments.

Similarly, higher scores on the experiences in the Systemic Awareness subscale may be expected in TA, but not in a more traditional assessment. This subscale asks parents to recognize a more systemic understanding of the child's problems including

how family struggles affect the child and that family members may also need to change in order to help the child. This is one of the intervention aspects of TA, but it may not by itself be highly related to overall parent satisfaction. Rather, it is conceptualized as a piece of TA that can contribute to significantly higher overall satisfaction ratings from the TA experience when compared to other assessment modalities.

Analysis

Descriptive Statistics

Basic descriptive statistics for demographics were reported. Chi-Square analysis was used to check for over/under representation of participants (gender, age, site, etc) with a significance level of .01 due to multiple analyses. Scale scores and items were evaluated for excessive skewness (>2.00) and kurtosis (>7.00) to determine whether measures were univariate normal (Curran, West, & Finch, 1996).

Scale Analysis

Research Question 1: What is the factor structure underlying the PEAS?

Confirmatory Factor Analysis (CFA) was used to evaluate the hypothesized six factor model of the PEAS. The initial model consisted of a first-order model with six factors, each representing a PEAS subscale (See Appendix B). The latent variables were allowed to correlate; correlations above .90 indicated that two subscales should be combined. In this way, the underling structure of the PEAS was tested by comparing nested competing models. Nested models are models that can be derived by constraining additional parameters in a model (Keith, 2006).

CFA allows for competing models to be evaluated by using fit statistics such as χ^2, root square mean of approximation (RMSEA), standardized root mean square residual

(SRMR), the comparative fit index (CFI), and Tucker-Lewis Index (TLI). For RMSEA, values below .08 indicated reasonable fit, and values below. 05 indicated good fit; for SRMR values below .08 also indicated good fit (Hu & Bentler, 1999). Values above .90 represented adequate fit, and values above .95 for CFI /TLI indicated good fit. Lower values of χ^2 in comparison to more degrees of freedom indicated a more parsimonious, and thus better, fit (Keith, 2006; Reynolds & Keith, 2009)

A hierarchal (bi-factor) CFA (see Appendix C) was conducted to determine factor loadings of the individual items on the subscales after controlling for general satisfaction. Past attempts to create multidimensional parent satisfaction scales have not controlled for overall parent satisfaction, which may have lead to a unidimensional result. A notable exception is research by Brannan et al. (1996) which included CSQ-8 items and created a distinct factor for general satisfaction. Although the PEAS is designed to measure dimensions distinct from general satisfaction, it was hypothesized that the items are also influenced by overall satisfaction. The process of testing competing models allowed for a more thorough exploration of the underlying structure of the PEAS than is generally undertaken for satisfaction questionnaires.

Once the best model, based on theory and statistical evidence, was demonstrated the results were used to refine the items on the subscales. To be retained, an item needed to have a minimum loading of .4, which is a common minimum used in item factor analysis (Brannan et al., 1996) and it was expected that most items will load above .60. Ideally, the final scale should contain no more than 40 items. Although this number is greater than many of the current satisfaction measures, the PEAS is also attempting to measure six subscales instead of only one or two factors for general satisfaction.

Psychometric properties including Cronbach's alpha for each subscale and corrected item-total correlations were also analyzed.

Research Question 2: Does the PEAS measure the same constructs across various groups?

To determine whether the PEAS items function similarly across groups, multiple group, mean and covariance structure analysis (MC-MACS) were used. Groups of interest included respondent type (parents vs. guardians), language (English vs. Spanish version), and social economic status (SES) as determined by education and income. After configural invariance was determined, successive levels of constraints, including metric, intercept, and residual invariance were tested. Difference in χ^2 and other fit statistics were used to determine if more parsimonious (constrained) models differed significantly from the configural model. If strong (metric and intercept) or strict (measurement) invariance was found, then group latent variances and means were compared for significant differences.

Data Analysis

Research Question 3: Which PEAS factors best predict overall satisfaction?

In order to aid program revision and service delivery, it is important to discover which interpersonal experiences are most predictive of overall parent satisfaction. To test this hypothesis, correlations with satisfaction were compared for significant relationships between subscales and general satisfaction. In addition, structural equation modeling was used to test a model showing the influence of each revised PEAS subscale on general satisfaction, as represented by the CSQ-8 scores (See Appendix D). It was expected that the relation between the parent and the assessor, the level of collaboration, and parent's

perception of the child-assessor relationship (as measured by the Parent-Assessor Relationship, Collaboration, and Child-Assessor Relationship subscales) would be the best predictors of overall parent satisfaction. The negative feelings associated with a parent needing a child assessment are often not included in parent satisfaction measures. It was hypothesized that lower levels of negative feelings will be associated with higher levels of parent satisfaction.

Research Question 4: Can the PEAS subscales differentiate between collaborative and traditional assessment practices?

Although the PEAS was expected to be related to general satisfaction, one of the major goals in creating the PEAS was to highlight the particular processes and outcomes hypothesized to promote change and understanding in TA. Although some of the PEAS factors were expected to score highly in all quality child assessments (particularly Child-Assessor Relationship and Parent-Assessor Relationship), the Collaboration, New Understanding of Child, and Systemic Awareness subscales were expected to be significantly higher in collaborative and Therapeutic Assessment.

The surveys collected from the various sites were grouped into 'traditional' and 'collaborative' types, based on whether sites used aspects of the Therapeutic Assessment process (see Finn, 2007; Tharinger, Krumholz, et al., in press). Independent t-tests were used to test for significant differences between assessment types on the PEAS subscale scores.

Chapter IV: Results

Database and IRB Approval

University IRB approval was received to create a database of pre-existing PEAS data for secondary analysis. Because the majority of data collection was already in place or completed through several student dissertation studies and various clinical and school sites, it was deemed that PEAS data would be released for secondary analysis for the current study. This procedure ensured that no further participation was required from participants, and helped maintain confidentiality since the data could be provided in de-identified form. Thus, a database was created in SPSS that pooled PEAS, CSQ, and demographic responses from non-identified participants at the various sites. All structural equation modeling was conducted using AMOS version 7.0.

Due to the multisite/multistudy nature of the data collection, there was some variability in the data collection procedures. The method of collection (paper, electronic, or phone) varied by each site's need and preference; however, the PEAS protocol was delivered consistently in terms of wording, order, and number of items.

Descriptive Statistics

Site Characteristics

A total of 138 PEAS protocols were considered for this study. Of those, 4 had been administered in Spanish. This small number of Spanish versions would not allow for a comparison between language versions, and therefore only English versions were used for this study. The following is a breakdown of the 134 remaining PEAS protocols by site: Site A, private neuropsychological clinic, 27% (n = 36); Site B, community clinic,

43% (n = 58); Site C, private assessment practice, 18% (n = 24); Site D, public school district, 12% (n = 16).

Site A is a private neuropsychological clinic that primarily evaluates learning disabilities, ADHD, autism spectrum disorders, traumatic brain injury, and other neurological disorders. Site A contributed both standard practice and collaborative assessment protocols to the database through a dissertation study evaluating the effects of adding a child feedback session and fable to the assessment process.

Site B is a community clinic that primarily serves at-risk families, including foster parents, adoptive parents, and legal guardians. Site B's assessment practice follows a collaborative/therapeutic model including gathering assessment questions, letters to parents, and child fables. The assessments at Site B primarily address emotional and behavioral problems, such as ADHD, anger problems, and externalizing behaviors.

Site C is a private assessment clinic that conducts assessment for learning disabilities, ADHD, autism spectrum disorders and independent educational evaluations. Site C is currently taking part in a dissertation evaluating the effects of adding TA components to their standard assessment practice. Only standard practice PEAS data was available for the current study.

Site D is a school district that conducts evaluations to determine eligibility for special education services, including learning disabilities and behavioral/emotional concerns. Site D contributed both standard practice and collaborative assessment PEAS protocols as part of a dissertation comparing the addition of components of Therapeutic Assessment.

Participant Characteristics

Respondents for the PEAS included biological parents (67%), adoptive/foster parents (17%), other kin guardians (10%), and non-kin guardians (5%). The majority of parent respondents were female (80%) rather than male (20%). The children and adolescents of the parents/guardians who received the assessments ranged from 4 to 18 years of age, with a mean of 9.72 years and standard deviation of 3.2 years. The majority of children and adolescents receiving assessments were male (64%) rather than female (36%). Child and adolescent descriptions of ethnicity included: 33% African-American, 28% Caucasian, 11% Hispanic, 2% Asian, 6% Biracial, and 19% not reported.

Chi-Square analysis demonstrated that there were no significant differences between sites in the proportion of respondent gender, χ^2 (3, n=134) = 3.50, p >.05, and child gender, χ^2 (2, n=105) = 1.81, p >.05. However, each site serves primarily different populations, and therefore it was not surprising that there were significant differences in proportions of child ethnicity and respondent education between sites. Sites A and C primarily serve higher income Caucasian families, while Sites B and C primarily serve lower income African American and Hispanic families.

Table 1: Site Characteristics

Characteristics	Site			
	A	**B**	**C**	**D**
Description	Private Neuropsychological Clinic	Community Clinic	Private Assessment Practice	Public School District
Collection Method	Paper & Electronic	Paper & Phone	Electronic	Paper
Primary Child Ethnicity	Caucasian	African American	Caucasian	Hispanic
General SES of Clients	High	Low	High	Low
Respondent Relationship to Child	Parents	Parents & Legal Guardians	Parents	Parents
n (PEAS)	36	58	24	16
n (CSQ)	36	55	24	0

General Satisfaction (CSQ)

A total of 115 CSQ scores were available for analysis, as one site did not collect parent satisfaction data. In addition, another site used either the CSQ-8 or a shortened version, the CSQ-4 (n = 23), to collect parent satisfaction. There were no significant differences in CSQ scores between the two versions for that site, $t(53) = 1.423$, $p > .05$. There were also no significant differences between the three sites that collected general satisfaction scores regardless of whether the CSQ-4 protocols were included, $F(2,114) = 1.373$, $p > .05$ or not $F(2, 91) = .869$, $p > .05$. Thus, all of the PEAS protocols were used for the scale revision, but only the 115 paired PEAS and CSQ protocols were used in later analyses investigating the relation of PEAS subscales to general parent satisfaction.

Items

Before conducting the pilot scale revision, each of the 64 items on the PEAS and 8 items on the CSQ were reviewed for skewness and kurtosis, which are indicators of

non-normality. Items 1, 34, 41, and 60 (The assessor worked well with my child, I felt the assessor was cold towards me, I felt the assessor looked down on me, and At the end of the assessment, I was left feeling angry) on the PEAS indicated non normality by having either skewness values greater than two or kurtosis values greater than seven (Curran, West, and Finch, 1996). These four items were noted as cautionary and their performance closely tracked in the PEAS scale revision. Fortunately, the vast majority of PEAS items (94%) met criteria for normality, indicating that analyses that assume univariate normality were appropriate.

In addition, 5 of the 8 CSQ items had substantial skewness or kurtosis (see Table 2). The negative skew for the CSQ items demonstrates the 'ceiling effect' of satisfaction scores, where item responses cluster between the highest two ratings. The CSQ uses a 4 point Likert Scale, and the average scores for the eight CSQ items ranged from 3.4 to 3.78, with standard deviations ranging from .47 to .66. The mode for all eight items was 4 and the median for seven of the eight CSQ items was also 4, which indicates that the 50th percentile of respondents was at the highest possible rating of the scale.

Table 2: Descriptive Statistics for Selected PEAS and CSQ Items ([a]Item has been reverse scored)

Item	Mean	SD	Skewness	Kurtosis
PEAS 1	4.57	.68	-2.02	6.04
PEAS 34[a]	4.49	.74	-2.18	6.97
PEAS 41[a]	4.56	.68	-2.14	6.86
PEAS 60	1.49	.78	2.44	7.86
CSQ 1	3.77	.53	-2.68	8.01
CSQ 2[a]	3.67	.61	-1.89	3.55
CSQ 3	3.40	.66	-.90	0.72
CSQ 4[a]	3.76	.61	-2.51	7.05
CSQ 5[a]	3.65	.65	-2.16	5.00
CSQ 6	3.53	.58	-.80	-.34
CSQ 7	3.70	.61	-2.46	7.23
CSQ 8[a]	3.78	.47	-2.04	3.53

Initial Scales

In addition to reviewing individual items, the pilot PEAS subscales and CSQ total score were also evaluated for normality and reliability. None of the PEAS subscales demonstrated substantial skewness or kurtosis. All of the pilot PEAS subscales had Cronbach alpha reliability ratings .75 or higher. The New Understanding of Child (α = .90), Parent Assessor Relationship (α = .86), and Child Assessor Relationship (α =.87) had the three highest reliabilities, while the Systemic Awareness (α =.80), Collaboration (α = .76), and Negative Feelings (α = .75) subscales had the three lowest reliabilities. These results indicate that the pilot version of the PEAS subscales are reliable scales and are adequate for investigations or studies using that version of the PEAS.

The CSQ scores demonstrated substantial negative skew (-2.12), which is not surprising given the non-normality reported in the individual CSQ items. In addition, both versions of the CSQ had Cronbach alpha reliability scores of .92, which is consistent with reliability estimates of the CSQ in previous studies.

Table 3: Reliability and Descriptive Statistics for Measures

Subscale	**Cronbach's Alpha**	**# of Items**	**Mean**	**SD**
New Understanding of Child	.90	14	3.78	.64
Parent Assessor Relationship	.86	10	4.30	.47
Collaboration	.76	10	4.08	.50
Child Assessor Relationship	.87	10	4.17	.59
Systemic Awareness	.80	10	2.94	.69
Negative Feelings	.75	10	1.80	.52
CSQ Total	.92	8	3.64	.50

Scale Analysis and Revision

Missing Data

The majority of the PEAS protocols were complete and no patterns of missing data were noted. However, it appears some respondents occasionally skipped an item on the PEAS when completing the form. Common, but not recommended, ways of handling missing data include pairwise or listwise deletion. More sophisticated methods of dealing with missing data are available with structural equation modeling techniques (Graham, 2009).

Although in most SEM applications, programs such as AMOS can use full information maximum likelihood estimation to account for missing data in the models, modification indices are only available when no missing data are present. Modification indices provide estimates of additional paths in a model that could improve the fit of a model. The use of modification indices was essential to the scale revision to help determine whether items should load on other subscales and allow for correlated unique error variances of the individual PEAS items. Therefore, maximum likelihood estimation was used to impute an SPSS database file that would account for missing data and allow for the calculation of modification indices in the subsequent analysis.

Model Testing

As mentioned previously, the models using Confirmatory Factor Analysis (CFA) discussed below will be nested and competing models, which allows for the use of comparative fit statistics. The change in chi square (χ^2) and degrees of freedom between nested models suggests whether an increase in fit (designated by a decrease in chi square) is 'worth' the loss of degrees of freedom (with higher values indicating the more desired

parsimony). The model adjustments below were conducted by determining whether a change, such as the addition of a path, produced a statistically significant better fit via change in chi square. The deletion of items does not provide nested models, and so fit was determined by decreasing values of AIC scores.

Change in chi square calculations were conducted in small increments, the majority after each modification (i.e., change of a single path or covariance) to ensure that the change in the model was appropriate. However, this resulted in more intermediate models between an initial and final model than can be reported. Instead fit statistics for the initial and final models, plus key intermediary models will be provided for the various models below.

First Order Model

The first model of the pilot PEAS tested was the first-order model showing the six latent factors (subscales) and the items designated to load on each subscale (See Appendix A). Before removing any items from the subscales, the initial model estimates and modification indices were used to see if any items should be cross loaded on an additional subscale. For example, the majority of the reverse-scored items initially loaded on the Negative Feelings subscale in addition to their primary subscales. Secondly, the unique errors of the PEAS items that had modification indices greater than thirteen were allowed to covary. This was the best fitting 'base' model of the pilot PEAS that included all of the items and potential item cross loadings, from which the scale would then be modified (See Appendix E)

Table 4: Fit Indices for PEAS Revision Models

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	*p*	TLI	CFI	RMSEA	SRMR	AIC
Initial	3852.539	1937								
Base	3405.014	1917	447.525	20	<.000	.66	.68	.076	.094	3859.014
Items below .5 removed (45 items)	1624.839	927	1780.175	990	<.000	.77	.78	.075	.094	1930.839
Subscales Reduced (27 items)	439.431	308	1185.408	619	<.000	.91	.92	.057	.078	633.431
5 Factor (25 items)	365.12	264	74.311	44	<.000	.92	.93	.054	.078	537.120

A cutoff of .5 was established for items to be retained on a subscale. Therefore, the first step of the revision was to remove paths or items that had loadings .3 or below. Because item loadings can change as the model is revised (as other items are removed), the model was re-estimated and then items with loadings of .49 and below were removed. Non significant correlations between subscales were also removed. Modification indices were then used to either allow items to load on an additional subscale if indicated or add further correlated unique errors for the remaining items. The resulting model consisted of 45 of the original 64 items.

The next phase of the revision involved reducing the number of items on each subscale. One of the goals for the revision was to reduce the PEAS to at most 40 items. Thus, although all items were loading above .5 on their primary subscale, the lowest loading items were subsequently removed. For the Child-Assessor Relationship subscale, item 1 was removed although it was the second highest loading item. This decision was based on the non-normality of item 1, mentioned previously. A similar decision was made for item 41, also removed due to its nonnormality. In addition, items with multiple correlated errors indicate that they may cross load on other subscales. However, unless

the path from the item to the additional subscale was high enough to be retained (> .5) on the subscale the path was removed from the model. Therefore, items with multiple correlated errors were eliminated in favor of items with less correlated errors so that the items retained on each subscale had a clear primary factor loading. After this phase of revision, each subscale had either four or five items and the PEAS now consisted of 27 of the original 64 items (See Appendix F).

Throughout the revision, the correlation between the Parent Assessor Relationship factor and the Collaboration factor remained high (.89 and above) and increased to .94 for the 27-item model. Due to this high correlation, a model combining these two factors was tested. Once the Parent Assessor Relationship and Collaboration subscales were combined, the three lowest loading items were removed so that the combined subscale consisted of seven items (See Appendix G). The five subscale model of the PEAS demonstrated better overall fit, noted by increases in TLI and CFI, decrease in RMSEA and decrease in AIC. The model testing steps of eliminating nonsignificant subscale correlations, checking for cross loading of items, and correlation of unique errors was repeated. Reliability statistics for the revised subscales was conducted at the conclusion of the revision and invariance testing process.

Second-Order Model

At this point in the revision process, there were still seven significant correlations among the latent factors. A second-order (hierarchical) model was tested to see if a general factor could account for the correlations among the subscales. When there is non-significant difference in change in chi square, the model with the higher degrees of freedom is preferred, in this case, the hierarchical model (See Appendix H).

The Parent-Assessor Relationship/Collaboration (PARC) factor had a loading of .96 on the general factor, which indicated that PARC was strongly associated with a general factor. Therefore, a model using the PARC subscale as the hierarchical factor did not result in statistically significant worse fit for the model, and with more degrees of freedom, became the preferred model. Through a review of modification indices, three correlated errors were added to the model that statistically improved fit. Throughout the PEAS revision, the fit statistics showed increasing evidence of better fit as noted through TLI and CFI increasing to approximately .95, RMSEA decreasing to below .05, SRMR below .08, and AIC decreasing with each model revision. The second-order model demonstrated adequate to good fit, and therefore is an appropriate model for the PEAS revision (See Appendix I).

Table 5: Fit Indices for Second Order Model Comparisons

Model	χ^2	***df***	$\Delta\chi^2$	Δdf	***p***	**TLI**	**CFI**	**RMSEA**	**SRMR**	**AIC**
First Order Five factor model	365.12	264				.92	.93	.054	.078	537.120
Second Order General model	366.228	266	1.108	2	0.574	.92	.93	.053	.079	534.228
Second Order PARC	366.638	268	.41	2	0.814	.93	.93	.053	.079	530.638
Second Order Best PARC	343.628	265	23.01	3	<.000	.94	.95	.047	.078	513.628

Bi-Factor Model

In addition to the second-order model, a bi-factor model was also analyzed. Rather than the hierarchical second-order model from the previous section, the bi-factor model controls for a general factor (satisfaction) by having each item load directly onto the general factor, rather than the subscales loading on the general factor (See Appendix

C). The bi-factor model was analyzed starting from the six subscale base model, with items loading on both the hypothesized PEAS subscales and the general factor. Thus, if item loadings continued to meet the .50 cutoff on the PEAS subscales, this indicated that the subscales were measuring a factor distinct from the general factor.

In the initial bi-factor model, multiple items on the Parent Assessor Relationship factor loaded significantly on the general factor (>.60), rather than the subscale (<.10). This indicated that the items on the parent assessor relationship scale could not be distinguished from the general factor. The results coincided with the previous finding from the hierarchical model, namely that the parent assessor relationship scale is so closely correlated to an overall factor, that it is indistinguishable. Thus, the Parent Assessor Relationship Factor was eliminated and its items loaded only on the general factor. Not surprisingly, a similar pattern emerged with the collaboration factor, with items not loading significantly on the subscale (<.20), but more highly on the general factor. This echoes the previous high correlation found between the Parent Assessor Relationship and Collaboration subscales, which were combined in the 5 subscale model. Thus, the items from the Parent Assessor Relationship and Collaboration subscales loaded only on the general factor in the bi-factor model. The same items retained previously on the Parent Assessor Relationship and Collaboration factor were retained as the highest loading items from those scales on the general factor in the bi-factor model.

The lowest loading items were then removed from the remaining factors, and the 25 items retained were the same as on the 5 subscale model. Items on the Child-Assessor Relationship subscale loaded almost evenly between the subscale factor and the general factor (approximately .50). The lowest item ("The assessor got my child to work really

hard") which had a loading of .40 was removed. Scale reliability was not affected by the removal of item 45. Two items remaining had loadings of .47 and .48, but could not be removed without eliminating the subscale. These results are similar to the previously high loading (.72) of Child Assessor Relationship on PARC in the second order model. Although distinct, the items which make up Child Assessor Relationship subscale load almost equally well on the general factor, but do not load as highly as the items from the Parent Assessor Relationship and Collaboration subscale.

Non-significant paths from the general factor to the 24 remaining items were eliminated, primarily on the Negative Feelings and Systemic Awareness subscales. This also echoed the previous findings, where Systemic Awareness was not significantly correlated to the second order PARC factor, and Negative Feelings had an inverse loading on PARC. The final bi-factor model consisted of 24 items, with the general factor the same Parent Assessor Relationship and Collaboration overarching factor found in the hierarchical model.

Table 6: Fit Indices for Bi-Factor Model Comparison

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	*p*	**TLI**	**CFI**	**RMSEA**	**SRMR**	**AIC**
Bi-Factor base model	3560.98	1884				.61	.64	.082	.994	4080.98
Bi-factor PARC on General only	2193.80	1235	1367.1	649	<.000	.70	.72	.076	.101	2583.80
Bi-factor 25 items	399.32	257	1794.5	978	<.000	.89	.90	.065	.0867	585.315
Bi-factor final	363.66	240	3.65	17	<.000	.90	.91	.062	.108	531.661
Second Order PARC	343.628	265	20.033	25	.745	.94	.95	.047	.078	513.628

The bi-factor model provides additional evidence that the PARC is the general factor measured by the PEAS, and that the items on the other subscales, even after loading on the general factor, combine to measure concepts distinct from the general factor. In accordance with the previous findings, New Understanding of Child had the second highest loadings on the general factor (.44), Child Assessor Relationship had the third (.38), Negative Feelings had two items load negatively, and Systemic Awareness had only one item significantly load on the general factor.

Overall, the bi-factor model did not reach the same level of fit as the second order PARC model previously tested. Even with one less item on the bi-factor model (which lowers AIC), the second order PARC has a lower AIC, greater TLI and CFI, and lower RMSEA. Additionally, the second order PARC model has a lower chi square, but greater parsimony (degrees of freedom), and has on overall higher level of fit. Even the best fitting bi-factor model, after deleting nonsignificant paths, has a greater AIC than the second order PARC model. Although the bi-factor model provides similar results, it appears the hierarchical second order model provides a better fit, both statistically and theoretically, for the PEAS.

Invariance Testing

Invariance testing through multiple group mean and covariance structures (MG-MACS) was conducted to ensure that the PEAS subscales were consistent and appropriate for use with multiple respondent types. The respondent relationship to the child data was available for all PEAS protocols. Initially, the respondent relationships were coded as biological, foster/adoptive, kin guardian (such as grandmother or aunt) and non-kin guardian. These groups were then recoded into two groups: biological parents (*n*

= 90) and non biological parents (n = 44) consisting of the foster, adoptive, kin, and non-kin guardians. In using confirmatory factor analysis for invariance testing groups should ideally be approximately the same size. For this analysis, the biological parent group was over twice the size of the nonbiological parent group. However, other breakdowns of the respondents into groups (e.g., gender, child gender) had even more disparate group sizes. Thus, the parent respondent type was used for this analysis.

Invariance testing requires that a separate model be produced for each group; with a sample size of 44 in one group, fit statistics using the full PEAS model of all six subscales would have less than adequate fit, due to the reliance of fit statistics on sample size. First, each subscale was tested individually for configural, metric, intercept and residual invariance. If strong (metric and intercept) invariance was found, then the factor variance and factor means were also tested for each subscale. Only after each subscale was tested was the invariance testing repeated for the full model.

The Parent Assessor Relationship and Collaboration subscale (PARC) was tested first and will be used as an example of the process repeated for each subscale (see Table 7). The configural model gives an overall fit when the sample is divided into the biological and nonbiological groups. The configural model should have very good fit, as it is the base model from which parameters will be increasingly constrained. Any changes, such as correlated errors, should be made at the configural model level. The configural model for the PARC had excellent fit, with TL and CFI both equal to one, RMSEA equal to zero, and AIC equal to 107.12. The configural model chi square and degrees of freedom consist of the sum of the χ^2 for each group separately. In the case of PARC, both the biological and nonbiological parent groups, when tested separately, had

fit statistics consistent with the configural model (TLI and CFI equal to one, RMSEA equal to zero).

To test for metric invariance, the factor loadings on the latent variable were set to be equal across groups. Metric variance provides evidence that the items are being interpreted similarly across groups, meaning that the items are measuring the same construct for both groups (Reynolds & Keith, 2009). Metric invariance is determined by the change in chi square in the same nested model process discussed previously for determining model fit. Metric invariance for PARC resulted in a nonsignficant difference in fit compared to the configural model. Therefore, metric invariance was established for PARC.

The next step for the invariance testing was to determine intercept invariance. Intercept invariance ensures that differences in observed means reflect differences in the latent means. Without intercept invariance, comparison of latent (and by extension, observed) means is not supported. The intercepts for each item were constrained across both groups for PARC and did not result in significantly worse fit. Thus, 'strong' invariance (both metric and intercept) was found for the PARC subscale, which indicates that the measurement instrument functions similarly across groups. Because the measurement aspects of the instrument were consistent across groups, latent variances and means could then be investigated.

Latent variances and means demonstrate differences across groups, whereas the previous steps to determine strong invariance were concerned with aspects of measurement across groups. Testing for factor variance provides information about whether the two groups have similar variability (Reynolds & Keith, 2009). Factor

invariance was held for PARC, indicating the variability for the responses of biological and nonbiological parents were the same. Lastly, invariance testing for factor means tests whether there are true differences of latent means between the groups. There were no significant differences in fit for factor mean invariance across groups, which indicate that there are not significant differences between biological and nonbiological parent scores on mean levels of PARC.

Table 7: Invariance Testing for Parent Assessor Relationship and Collaboration

Model	χ^2	***df***	$\Delta\chi^2$	Δdf	***p***	**TLI**	**CFI**	**RMSEA**	**AIC**
Bio parent only	8.016	13				1.027	1.00	.00	52.016
Nonbio parent	11.050	13				1.028	1.00	.00	55.050
Configural	19.124	26				1.027	1.00	.00	107.124
Metric	25.480	32	6.356	6	0.385	1.021	1.00	.00	101.48
Intercept	28.209	38	2.729	6	0.842	1.027	1.00	.00	92.209
Factor Variance	28.284	39	0.075	1	0.784	1.028	1.00	.00	90.284
Factor Mean	29.672	40	1.388	1	0.239	1.027	1.00	.00	89.672

Overall, the PARC demonstrated strong invariance for the biological and nonbiological parent groups. The change in chi-square between nested models was not significant, TLI and CFI remained high, and AIC consistently decreased, indicating better fit as models were increasingly constrained. This indicates that in terms of measurement, the PARC subscale is measuring the same construct across both groups. In addition, the variance and mean scores between biological and nonbiological parents are not significantly different.

The same process of invariance testing was repeated for the other four PEAS subscales. Strong factorial invariance was found for the Systemic Awareness and Negative Feelings subscales. No differences in factor variability or latent means were

found between groups for Systemic Awareness. Although there was not a significant difference in variability for the Negative Feelings subscale, there was a statistically significant difference between factor means. The degradation in fit across the fit statistics underscores that factor mean invariance did not hold for the Negative Feelings subscale: TLI and CFI drop to below .95, RMSEA is above .05, and AIC increases. Nonbiological parents had an average score .23 points higher than biological parents on the Negative Feelings subscale. This indicates that nonbiological parents report significantly higher levels of negative feelings than biological parents as measured by the PEAS.

Table 8: Invariance Testing for Systemic Awareness

Model	χ^2	***df***	$\Delta\chi^2$	Δdf	***p***	**TLI**	**CFI**	**RMSEA**	**AIC**
Configural	0.342	2				1.06	1.00	.00	52.342
Metric	3.125	5	2.783	3	0.426	1.03	1.00	.00	49.125
Intercept	7.959	8	4.834	3	0.184	1.00	1.00	.00	47.959
Factor Variance	18.794	14	2.693	1	0.100	0.97	0.97	.051	46.794
Factor Mean	18.797	15	0.003	1	0.956	0.98	0.98	.044	44.797

Table 9: Invariance Testing for Negative Feelings

Model	χ^2	***df***	$\Delta\chi^2$	Δdf	***p***	**TLI**	**CFI**	**RMSEA**	**AIC**
Configural	5.16	4				.97	.99	.047	53.16
Metric	8.642	7	3.482	3	0.323	.98	.99	.042	50.642
Intercept	11.116	10	2.474	3	0.480	.99	.99	.029	47.116
Factor Variance	13.839	11	2.723	1	0.099	.98	.98	.044	47.839
Factor Mean	18.452	12	4.613	1	0.032*	.95	.95	.064	50.452

The New Understanding of Child subscale met the fit requirements for metric invariance and partial intercept invariance. Full intercept invariance did not hold for the New Understanding of Child scale, and resulted in a significantly worse fit ($p < .05$). The TLI and CFI fit indices dropped to .96, RMSEA became .74 and AIC increased to 75.182. When intercept invariance does not hold, individual intercepts can be released to

determine if partial intercept invariance is a possibility. In the case of New Understanding of Child, intercept invariance was not significant when all items except item 58 (I have lots of new ideas about how to parent my child) were constrained to be equal. This indicates that item 58 has different intercepts for biological (3.33) versus nonbiological parents (3.74): in this case, nonbiological parents have a higher intercept. A higher intercept indicates that the item for nonbiological parents starts at a higher point, given the same latent factor mean.

Although full intercept invariance would be ideal, the partial intercept invariance, particularly for this item makes sense. Biological parents have generally been parenting their children for their child's entire lives, whereas nonbiological parents may have had much shorter time and experience parenting their children, particularly foster parents or guardians. Hence, nonbiological parents may have a much higher starting point in scores, because they have more to learn about parenting techniques for a specific child than biological parents.

Further testing of factor variance and factor means for the New Understanding of Child subscale found a significant difference in factor variance for biological and nonbiological parents, but no significant differences in factor means.

Table 10: Invariance Testing for New Understanding of Child

Model	χ^2	***df***	$\Delta\chi^2$	Δdf	***p***	**TLI**	**CFI**	**RMSEA**	**AIC**
Configural	9.142	10				1.01	1.00	.000	69.142
Metric	11.441	14	2.299	4	.681	1.01	1.00	.000	63.441
Intercept	31.182	18	19.741	4	.001*	0.96	0.96	.740	75.182
Partial Intercept	18.91	17	7.469	3	.058	0.99	0.99	.029	64.91
Factor Variance	24.80	18	5.89	1	.015*	0.98	0.98	.054	68.84
Factor Means	18.915	18	.005	1	.944	1.00	1.00	.020	62.915

Lastly, the Child Assessor Relationship (CAR) subscale was tested for metric and intercept invariance. The initial configural model had much better fit for the nonbiological parents than the biological parents. This indicated that the 'base' model was not a good starting point for comparisons, although the overall fit of the model was acceptable. Upon investigation, item 45 (The assessor got my child to work really hard) had very different factor loadings for each group. This was the same item that had been removed in the bi-factor model due to a low loading on the subscale after controlling for the general factor. For biological parents, the unstandardized factor loading was .62, the only standardized loading below 1.00 and the lowest item on the subscale. For nonbiological parents, the unstandardized factor loading was 1.95, the highest of all the items on the subscale. This misfit continued when metric invariance was tested; it produced significantly worse fit than the configural model.

A revised configural model was created by eliminating item 45, and the fit was very good for both groups. Although item 45 had loaded highly on CAR for the general scale revision, it appeared to be the only item for any subscale that resulted in metric invariance not being met. Item 45 was clearly functioning quite differently for biological and nonbiological parents, and was removed from the CAR subscale. Other items from the original CAR subscale which had previously been eliminated due to lower factor loadings were reconsidered to see if a replacement item was necessary. However, the items which had been previously eliminated showed the same difficulty as item 45 in terms of fit. Thus, the CAR subscale consists of the four remaining items from the initial configural model.

The revised configural model had very good fit and metric invariance held (see Table 11). In the intercept testing, partial intercept invariance was found. Item 49 (My child and the assessor really connected well) had an intercept of 3.78 for biological parents and 4.29 for nonbiological parents. Thus, it appears that although item 45 measures the same construct across the parent groups, nonbiological parents have a much higher starting point, a full half point higher on the 5 point rating scale, than biological parents. There were no factor variance or mean differences found between the parent groups with further testing, however, the testing was conducted with only partial intercept invariance.

Table 11: Invariance Testing for Child Assessor Relationship

Model	χ^2	***df***	$\Delta\chi^2$	Δdf	***p***	**TLI**	**CFI**	**RMSEA**	**AIC**
Configural	3.479	4				1.01	1.00	.000	51.479
Metric	4.584	7	1.105	3	.776	1.02	1.00	.000	46.584
Intercept	18.146	10	13.562	3	.004*	0.95	0.96	.079	54.146
Partial Intercept	7.339	9	2.755	2	.252	1.02	1.00	.000	45.339
Factor Variance	7.75	10	0.411	1	.521	1.01	1.00	.000	43.750
Factor Means	8.904	11	1.154	1	.283	1.01	1.00	.000	42.904

Residual invariance was also tested for each subscale. Residual invariance, which can be established after metric and intercept invariance, helps assure that there is no bias in the measurement instrument: group membership does not matter. Meeting residual invariance is deemed 'strict' invariance, but is not necessary to test factor means and variance. Only 'strong' (metric and intercept) invariance is needed to test for structural elements of a measure, such as factor variance and mean comparisons. Residual invariance was not found for any PEAS subscale except the Systemic Awareness subscale. The Systemic Awareness subscale is the most family oriented subscale in terms

of determining how much parents see the wider family system as contributing to or maintaining a child's difficulties. It is therefore interesting to note that the only PEAS subscale to hold residual invariance is the one that one might think would be most affected by being in a different parent groups. However, the Systemic Awareness subscale is unbiased towards either biological or nonbiological parents: group membership does not influence latent or measured mean scores. This finding places the Systemic Awareness scale at an even higher standard of psychometric strength and demonstrates that it is appropriate to use with various parent respondent groups.

Once each subscale had been tested for invariance, the full model of the 24 item revised PEAS was tested. Findings of the full model were consistent with the previous findings for each subscale. Metric invariance held, as did partial invariance by freeing the two items noted previously on the New Understanding of Child and Child Assessment Relationship subscales. Partial residual invariance was held for Systemic Awareness and New Understanding of child; all but two items on the Parent Assessor Relationship and Collaboration subscale met residual invariance. Residual invariance did not hold for Child Assessor Relationship, and only one item on the Negative Feelings subscale met residual invariance. Six of the seven correlated errors among test items were invariant across groups.

After the measurement aspects of the PEAS had been tested via metric and partial intercept invariance, structural aspects and differences between groups could be compared. The factor loadings of the subscales on Parent Assessor Relationship and Collaboration (PARC) and first-order factor covariances were invariant across groups, with the exception of the loading from Negative Feelings onto PARC. Negative Feelings

has an unstandardized factor loading of -.51 for biological parents and -1.27 for nonbiological parents. This indicates that the loading of Negative Feelings on PARC is significantly greater for nonbiological than biological parents. Next, first order latent intercepts, as well as the factor variance and mean for PARC, were tested for invariance. All loadings were invariant with the exception of Negative Feelings. Consistent with the individual subscale findings, nonbiological parents had a starting point .30 points greater than biological parents on Negative Feelings. Lastly, the unique variance of the first order factors was held across both groups for all subscales except New Understanding of Child. Congruent with the subscale invariance testing, the scores of nonbiological parents had greater variance on New Understanding of Child than biological parents.

Overall, the maintenance of fit across the increasingly constrained models of the hierarchical PEAS is demonstrated by nonsignificant changes in chi square, consistent TLI, CFI, and RMSEA values, as well as decreasing AIC. The full model invariance findings were consistent with those of the individual subtests, with the additional finding that Negative Feelings has a significantly greater factor loading on Parent Assessor Relationship and Collaboration for nonbiological than biological parent groups. Although residual variance is not required to compare latent means and variances, the full model demonstrated that the Systemic Awareness and New Understanding of Child subscales had residual invariance, as well as five out of seven items on PARC.

Table 12: Invariance Testing for Second Order Model of Revised PEAS

Model	χ^2	*df*	$\Delta\chi^2$	Δdf	*p*	TLI	CFI	RMSEA	AIC
Configural	584.433	482				.92	.93	.040	916.433
First Order Metric	608.839	501	24.406	19	0.181	.92	.93	.040	902.839
First Order Partial Intercept	630.456	518	21.617	17	0.200	.92	.93	.041	890.456
First Order Partial Residual	652.087	539	21.631	21	0.421	.92	.93	.040	870.087
Second Order Factor Loadings	659.862	543	7.775	4	0.100	.92	.92	.040	869.862
First Order Latent Intercepts	664.457	548	4.595	5	0.467	.92	.92	.040	864.457
First Order Factor Unique	668.177	551	3.72	3	0.293	.92	.92	.040	862.177

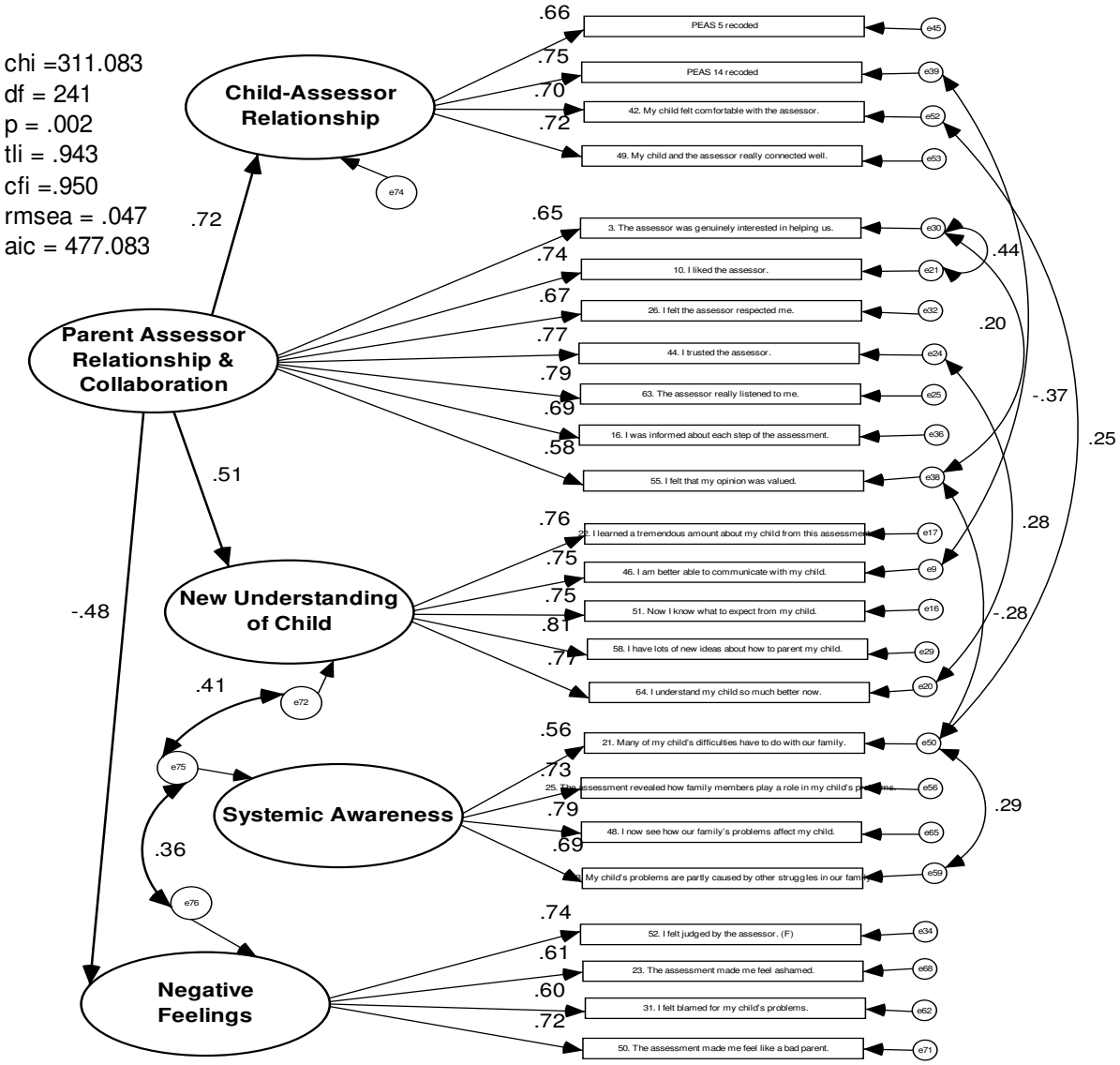
Final Model Interpretation

Figure 2 shows the Final Second Order Model of the PEAS revision, reorganized for easier interpretation. The fit for the final model is shown in Table 13. The revised PEAS had some significant differences from the pilot version that are worth noting. The first and most obvious change in the revised PEAS is the reduction from six subscales to five subscales. During the revision, the Parent Assessor Relationship and Collaboration subscales consistently demonstrated a correlation greater than .90. When these two subscales were combined into one subscale, the fit for the model improved significantly. When reviewing the process of finding the most highly loading items for the Collaboration subscale, qualitatively, the original items dealing with the teamwork present in Therapeutic Assessment (helping make sense of the test results, working as a

Table 13: Fit Indices for Revised PEAS

Model	χ^2	*df*	TLI	CFI	RMSEA	SRMR	AIC
Revised PEAS 24 Items	311.083	241	.94	.95	.047	.079	477.083

Figure 2: Revised PEAS Model



team, having a say in what the assessment focused on, being asked if the findings seemed right) did not load as highly as the items retained. It appears that the items that did load on the Collaboration subscale dealt more with being well informed about the assessment (being informed about each step of the assessment, understanding the goals, and having parents' opinion valued). It may be that the more TA based collaboration items were not

retained because they were not as widely applicable to the respondents from the various sites. For example, common assessment batteries may have been used, which do not allow for as much input from the parents. Or, parents may not expect to be consulted about the meaning of the results, as that is the assessor's job as the 'expert', but still want to be informed about the process and goals of the assessment. Overall, the two items from the Collaboration subscale that were retained on the combined Parent-Assessor Relationship/Collaboration subscale address being informed about the assessment and feeling the parent's opinion is valued. These items qualitatively coincide with the items from the original Parent-Assessor Relationship subscale, such as the assessor really listening to the parent, feeling respected, the assessor as genuine, and trusting the assessor.

The second major pattern in the revised version of the PEAS is that the reverse scored items were not retained on any subscale, with the exception of items 5 and 14 on Child Assessor Relationship. Rather, the negatively worded items all cross-loaded on the negative feelings subscale. The Negative Feelings subscale in the pilot version included items that specifically implied negative feelings about the assessment. However, one reverse scored item originally designed for another subscale (I felt judged by the assessor for the Assessor-Parent Relationship subscale) loaded highly enough on the negative feelings subscale to be retained. Qualitatively, it is not surprising that feeling negatively evaluated by the assessor would be related to the other three items on the subscale describing feeling ashamed and like a bad parent. Also, although negative or reverse scores items are used in scales to avoid response bias, they often function differently than their positively worded counterparts (DeVellis, 2003). Thus, the Negative Feelings

subscale remained distinct from the other subscales and incorporates negative feelings evoked by both the assessor and the assessment process.

Although the PEAS revision initially started as a first order model, the final model shows a hierarchical relation between the Parent-Assessor Relationship and Collaboration (PARC) subscale and the other subscales. The strongest relationship with the PARC subscale is the Child-Assessor Relationship subscale with a loading of .72. The Negative Feelings subscale is inversely related to the PARC subscale (-.48), which indicates that as the parent assessor relationship increases, the negative feelings about the assessor and the assessment decrease. The New Understanding of Child subscale has a .51 loading on the PARC subscale.

The Systemic Awareness (SA) subscale demonstrates the most unique relation to the other subscales in the revised PEAS model. Noticeably, the SA subscale did not significantly load onto the PARC subscale. Rather, the SA subscale is significantly correlated with New Understanding of Child (.41) and Negative Feelings (.36). Rather than PARC having a direct effect on SA, it appears it may indirectly affect Systemic Awareness by decreasing Negative Feelings and increasing a New Understanding of Child.

Another important aspect of the SA subscale is that it is positively correlated with the Negative Feelings subscale, indicating that increases in systemic awareness scores are associated with increases in negative feelings. One hypothesis for this finding is that acknowledging the family's role in a child's problems leads to negative feelings; increases in SA may be akin to having to admit some culpability or 'guilt' on the part of

the family. The implications of this relationship in the context of the theory of TA will be explored more thoroughly in the discussion section.

Additional Research Questions

Relationship of PEAS Subscales to Satisfaction

Two-tailed Pearson Correlations were computed between CSQ scores, the pilot PEAS subscales, and the revised PEAS subscales.

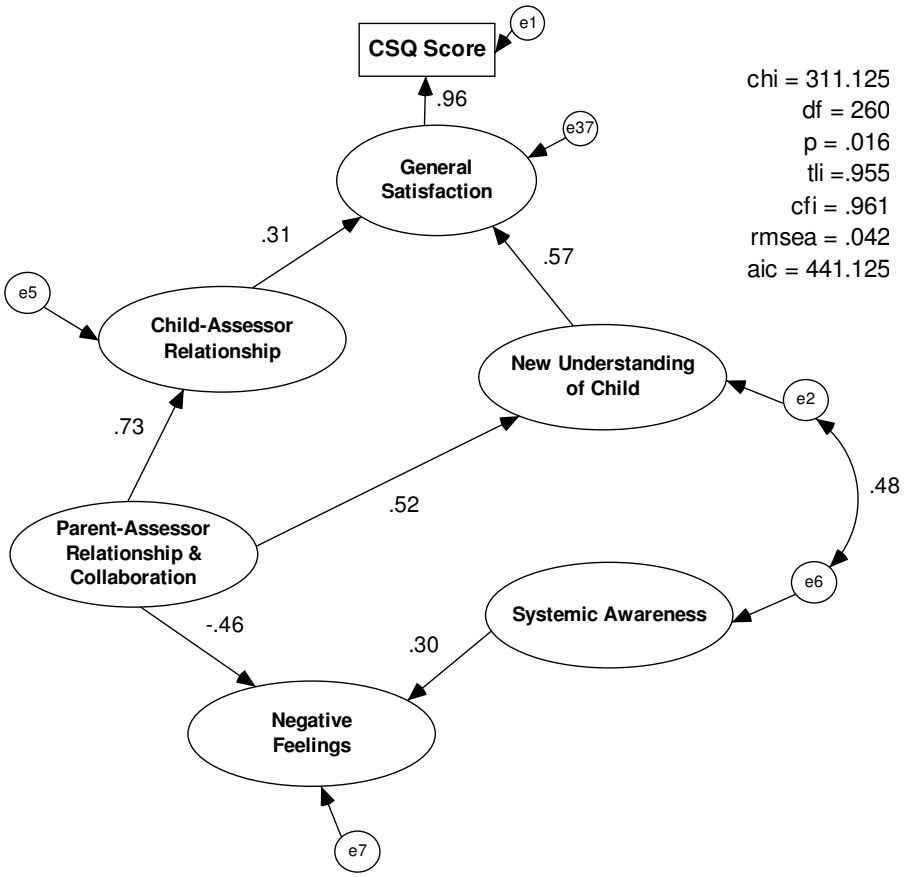
Table 14: Correlation of PEAS subscales with CSQ score (** significant at .01; * significant at .05)

Subscale	Correlation with CSQ	
	Pilot	Revised
New Understanding of Child	.70**	.64**
Parent Assessor Relationship	.54**	.48**
Collaboration	.66**	-
Child Assessor Relationship	.56**	.45**
Systemic Awareness	.15	.21*
Negative Feelings	-.33**	-.24**

In the pilot version of the PEAS, all of the subscales were significantly correlated with satisfaction with the exception of Systemic Awareness. The revised PEAS subscales maintain significant correlations with CSQ, and Systemic Awareness now has a significant correlation with CSQ at the .05 level. The correlations between the PEAS subscales and CSQ decreased slightly between the Pilot and Revised versions, however, this is most likely due to the reduced length of the subscales. Notably, the PEAS revision was conducted independently of an item's relationship to CSQ, so there may have been items on the pilot PEAS version that may have been more strongly correlated to satisfaction, but were not retained because they did not load as highly on the individual subscales. When analyzed independently, each of the revised subscales is significantly correlated to CSQ. However, further analysis of the relationships and loadings among the subscales themselves were also conducted.

Structural Equation Modeling was used to investigate the latent factors of the PEAS subscales (using the revised version) and their relationship to general satisfaction, as represented by CSQ scores. Only 115 CSQ were available in the data set, and so this model uses an *n* of 115, instead of the 134 used for the PEAS revision. Figure 3 shows the same relationships between the PEAS subscales as given in the final second order model of the PEAS revision, including correlations among error variances.

Figure 3: Structural Equation Modeling of PEAS and General Satisfaction



*Shown without item indicators for PEAS subscales

The magnitude and direction of the subscale relationships were consistent with the PEAS revision model. PARC is positively related to Child Assessor Relationship (.73) and New Understanding of Child (.52) and negatively related to Negative Feelings (-.46). Both Negative Feelings (.30) and New Understanding of Child (.48) are significantly related to Systemic Awareness. Because PARC is the hierarchical latent factor in the revision, it is depicted as directly influencing the three subscales it is connected too, which is why a directional arrow is used. However, the relationships between Systemic Awareness and New Understanding are correlations because the directionality is unknown.

The CSQ scores are represented by a rectangle with a loading of .96 onto the latent General Satisfaction variable. When using a single measured indicator for a latent variable, setting the error variance can account more accurately for error. The error variance for the CSQ was derived from the reliability and variance of the measure and set to .0204. Initially, a path from each subscale was drawn to General Satisfaction, however, only the New Understanding of Child (.57) and Child Assessor Relationship (.31) subscales had significant direct effects on CSQ. It appears that PARC indirectly affects (.521) General Satisfaction through increased New Understanding of Child and Child Assessor Relationship. Tables 15 and 16 provide a listing of direct, indirect, and total effects for the model.

Table 15: Standardized Direct, Indirect, and Total Effects on General Satisfaction

Subscale	Direct	Indirect	Total
New Understanding of Child	.565	-	.565
Parent Assessor Relationship and Collaboration (PARC)	-	.521	.521
Child Assessor Relationship	.315	-	.315

Table 16: Standardized Direct, Indirect, and Total Effects Between Subscales

Subscale	Direct	Indirect	Total
PARC on Child Assessor Relationship	.728	-	.728
PARC on New Understanding of Child	.517	-	.517
PARC on Negative Feelings	-.457	-	-.457
Systemic Awareness on Negative Feelings	.302	-	.302

Differentiation between Assessment Types

Each assessment was designated as either Traditional/Standard Practice (n = 55) or Collaborative/Therapeutic Practice (n = 79). Traditional or Standard Practice assessments were those assessments that followed the standard assessment procedure for that site, without any additional elements or training from Therapeutic Assessment. The Collaborative/Therapeutic assessments were those which directly incorporated a piece of TA, such as child fables or assessment questions, although none used a comprehensive model (See Table 17). Independent t-tests were conducted to test for significant differences between assessment types. The Collaborative/Therapeutic assessments had greater average scores for New Understanding of Child, Parent Assessor Relationship and Collaboration, Child Assessor Relationship, and Negative Feelings. Only the Systemic Awareness subscale had a statistically significant difference across groups (see Table 18). These results indicate that Collaborative/Therapeutic assessments have significantly higher levels of average Systemic Awareness than the Traditional/Standard Practice Assessment. The higher levels of Negative Feelings is also greater for Collaborative/Therapeutic assessments, but was not significant for this study. As mentioned previously, the PEAS revision indicated that as Systemic Awareness

increases, so does Negative Feelings, and this has been replicated based on assessment type.

Table 17: Descriptive Statistics for PEAS Subscales by Assessment Type

Subscale and Group	Mean	SD
New Understanding of Child		
Collaborative/Therapeutic[a]	3.78	.856
Traditional[b]	3.68	.642
Parent Assessor Relationship and Collaboration		
Collaborative/Therapeutic	4.42	.499
Traditional	4.29	.562
Child Assessor Relationship		
Collaborative/Therapeutic	4.24	.676
Traditional	4.11	.707
Systemic Awareness		
Collaborative/Therapeutic	2.90	.978
Traditional	2.54	.839
Negative Feelings		
Collaborative/Therapeutic	1.72	.704
Traditional	1.50	.535

[a]n = 79 for each subtest. [b]n = 55 for each subtest

Table 18: Independent *t*-test results for PEAS Subscale by Assessment Type

Subscale	*t*	*df*	*p*	*d*
New Understanding of Child	.743	132	.459	.13
Parent Assessor Relationship and Collaboration	1.441	132	.152	.25
Child Assessor Relationship	1.023	132	.308	.19
Systemic Awareness	2.241	132	.027*	.40
Negative Feelings	1.902	132	.059	.35

Demographic Findings

Various demographics for child and parent respondents were compared via ANOVA or independent t-tests to check for any significant differences. The complete statistical results for each category are provided in Appendix K. Child ethnicity was recoded into four groups: Caucasian (n = 37), African American (n = 45), Hispanic (n = 15), and Other (n = 12). One way ANOVA found no significant differences for any of the PEAS subscales or CSQ scores between groups of child ethnicity. Independent t-tests found no significant differences on average PEAS and CSQ scores for female (n = 107) and male (n = 27) parent respondents. In addition, comparison of scores based on child gender (female: n = 38; male: n = 67) found no significant differences for PEAS or CSQ scores. These results were consistent with findings in previous satisfaction research that found demographic variables did not significantly predict satisfaction scores.

Reliability and Scoring Guidelines

Cronbach alpha reliability and descriptive statistics for the five revised PEAS subscales were calculated and are listed below. See Appendix L for full tables of scale and item-scale statistics.

Table 19: Reliability and Descriptive Statistics for Revised PEAS Subscales

Subscale	Alpha	# Items	Mean	SD	Skewness	Kurtosis
Parent Assessor Relationship and Collaboration	.88	7	4.37	.53	-0.678	0.378
New Understanding of Child	.88	5	3.74	.77	-0.983	1.810
Child Assessor Relationship	.79	4	4.17	.69	-1.080	1.933
Systemic Awareness	.80	4	2.75	.94	0.045	-0.572
Negative Feelings	.76	4	1.63	.65	1.113	1.878

The total points possible for each subscale, as well as mean sum of subscale, standard deviation, and 20th and 80th percentile scores are shown in Table 20. All of the revised PEAS subscale scores have nonsignificant skewness and kurtosis scores, indicating that they do not suffer from the 'ceiling effect' so often cited in the satisfaction literature.

In order to facilitate interpretation of the PEAS subscale scores, the Total Score (sum of Likert Ratings) for each subscale has been converted to T-Scores. T-Scores have a mean of 50 and a standard deviation of 10; T-scores between 40 and 60 are considered average. Table 21 provides general scoring guidelines for Total Scores, while Appendix M provides a conversion table from subscale total scores to t-scores.

Table 20: Descriptive Statistics and Percentiles for PEAS Subscales

Subscale	Total Points	Mean	S.D.	20th	80th
Parent Assessor Relationship and Collaboration	35	30.57	3.70	27	34
Child Assessor Relationship	20	16.74	2.75	15	19
New Understanding of Child	25	18.72	3.87	16	22
Systemic Awareness	20	11.00	3.75	8	15
Negative Feelings	20	6.52	2.56	4	8

Table 21: Scoring Guidelines for revised PEAS Subscales

Subscale	T-Score		
	≤40	41 - 59	≥60
Parent Assessor Relationship and Collaboration	≤ 27	28-33	35
Child Assessor Relationship	≤ 14	16-18	20
New Understanding of Child	≤ 15	17-22	≥ 23
Systemic Awareness	≤ 7	9-14	≥ 15
Negative Feelings	4	5-8	≥ 9

Each site will receive individual feedback consisting of a summary of the findings, overall PEAS subscale and CSQ scores, as well as the site's own scores, broken down by assessor if requested. Thus, the PEAS scoring guidelines will allow sites to compare their overall and even individual assessor performance to help further understand how parents experienced their child assessment process.

Chapter V: Discussion

Findings for Research Questions and Hypotheses

Research Question 1

The first research question in this study concerns the underlying factor structure of the revised PEAS based on the pilot sample. It was hypothesized that a first order six factor model would be found; however, a second order five factor model was determined to be the best fitting model. The second order model implies that the hierarchical latent variable (Parent Assessor Relationship and Collaboration; PARC) is a general factor that influences the other subscales. The Child-Assessor Relationship, New Understanding of Child, and Negative Feelings subscales load significantly on the hierarchical PARC. It was expected that the PEAS subscales would be correlated, however, correlations do not provide directionality or causality. The hierarchical model provides an implication of directionality, that the correlations among the other subscales can be explained by a general factor that influences the scores on the other subscales. The finding that the PARC subscale loaded so highly when a hierarchical general factor model was tested, so as to be almost indistinguishable (.96), indicates that the interpersonal relationships between the parent and assessor is a overarching construct influencing the other subscale/processes measured by the PEAS. This finding was also replicated with the bi-factor model, where items from the Parent Assessor Relationship and Collaboration subscales loaded so highly on the general factor, that those subscales were eliminated and subsumed by the general factor.

The finding that the PARC subscale is measuring an overall, hierarchical factor as related to the other PEAS subscales is supported by Therapeutic Assessment (TA) theory and intent as well. In TA, the relationship between the assessor and client, in this case parents, is regarded as fundamental to the success and potential impact of the assessment. The assessor strives to create a safe holding environment for the client that will then allow the client to explore and incorporate the assessment results more fully. The PARC subscale was designed to measure the interpersonal relationship between the assessor and parent, and the CFA model has gone a step further by showing that the interpersonal relationship is an overarching construct that influences three of the other subscales.

In addition to the second order model of the PEAS revision, the revised version includes only five subscales rather than the hypothesized six. As discussed in the results section, the Collaboration subscale correlated above .90 with the Parent Assessor Relationship subscale, and their combination provided better, more parsimonious fit for the model. The original Collaboration items were a mixture of items reflecting that parents were informed and involved in the assessment. The items most specific to Therapeutic Assessment (i.e. I helped make sense of the test results, the assessor asked me if the findings seemed right to me, I felt like part of a team working to help my child) may not have been applicable in the current sites.

Although the majority of the assessments that comprise this sample were collaborative or therapeutic in nature, real world time and budget constraints did not allow sites to follow the Comprehensive TA model for children. In fact, the goal for many of the studies included in the database was to see if adding a selected piece or pieces of Therapeutic Assessment could make a significant difference in standard

practice or traditional assessment. Thus, items more specific in nature to the parent experience during a comprehensive TA, such as parents viewing and processing the assessment sessions with an assessor, may not have been as relevant across different types of settings and assessments, and so were not as feasible items to be retained in the revision of the PEAS. Rather, the items that loaded the most highly on Collaboration (before it was combined with PAR) include understanding the goals of the assessment, being informed about the assessment process, and feeling the respondent's opinion was valued. Therefore, the combination of the Collaboration and Parent Assessor Relationship subscales reflects aspects of a solid interpersonal relationship between assessor and parent; items let parents communicate how heard, valued, informed, respected, genuine, and trustworthy they found their experience with the assessor.

The other four subscales hypothesized on the pilot PEAS were maintained in the revised model. The New Understanding of Child, Child Assessor Relationship, Systemic Awareness, and Negative Feelings subscales have been reduced from the pilot version to the four or five highest loading items per subscale. During the revision process some items may have cross loaded on other scales, but these cross loadings were below the minimum .5 cutoff, and so were not retained. The exception is the item which was originally on the Parent-Assessor Relationship subscale that loaded more strongly on the Negative Feelings subscale. Thus, the Negative Feelings subscale, originally designed to measure negative feelings about the assessment process now includes negative feelings about being judged by the assessor as well.

Overall, the hypothesized first order six factor model of the PEAS was not found to be the best fitting model, and the revised version consists of a second order five factor

model. The revised model has adequate fit statistics and helps theory take a step forward by showing the fundamental/hierarchical nature of the interpersonal relationship between the parent and the assessor.

Research Question 2

The second research question for this study was to compare the revised PEAS across various groups to ensure that the scale measures the same constructs consistently across the groups. Unfortunately, there were not enough Spanish PEAS protocols to allow for a language comparison, and so only English protocols were used for the current revision. However, a Spanish translation can be created for the revised PEAS and a test of invariance across language versions should be a priority for future studies. In addition, respondent education was not consistently reported within the database, and income data was unavailable for the PEAS respondents. Although general estimations about the SES of typical clients served by each site can be made, no consistent data for an analysis at this level was available.

Invariance testing was conducted for different parent respondent groups: biological and nonbiological. One cautionary note is that the sample sizes were not as even as is typically suggested for invariance testing; the biological parent group was twice the size of the nonbiological parent group. The nonbiological parent group consisted of adoptive, foster, kin, and non-kin guardians. The invariance testing led to the removal of item 45 on the Child Assessor Relationship subscale due to its different performance across the respondent groups. This item had also been removed in the bi-factor model, and was viewed with caution upon beginning the invariance testing. All subscales met metric invariance, indicating that the same constructs are being measured

across the parent respondent groups. Three of the subscales (Parent Assessor Relationship and Collaboration, Systemic Awareness, and Negative Feelings) had full intercept invariance, which met criteria for 'strong' factorial invariance. The New Understanding of Child and Child Assessor Relationship subscales met partial intercept invariance, meaning one item on each subscale had a different intercept, or starting point, for nonbiological parents than biological parents. Overall, it appears that the PEAS subscales are measuring the same constructs across parent respondent groups, and hence the measurement aspects of the revised PEAS are sufficient to compare latent variance and means.

Both factor variances and means were invariant across parent groups for all PEAS subscales with two exceptions. The New Understanding of Child had different variances, or different sizes of the normal curve, for biological and nonbiological parent groups. In addition, a significant mean difference was found for the Negative Feelings subscale, with nonbiological parents reporting significantly higher levels of Negative Feelings about the assessment. Lastly, the Systemic Awareness subscale met criteria for 'strict' factorial invariance with the addition of residual invariance. This indicates that the Systemic Awareness subscale is unbiased; group membership does not influence SA scores and this indicates a high level a psychometric properties.

After testing for measurement and structural invariance for each subscale, the full hierarchical model was tested as well. The results were consistent with the individual subscale results, with the addition of a significant difference in factor loading of Negative Feelings on PARC. The Negative Feeling scores of nonbiological parents had a significantly greater inverse loading on PARC than the scores of biological parents.

In conclusion, invariance testing was conducted on the PEAS subscales to increase the level of psychometric rigor of the scale. Often, measured means are compared (such as in ANOVA testing) without the scale having been tested for invariance. Invariance provides statistical evidence that the scale is measuring the same constructs across groups (metric invariance) and that the construct functions similarly (intercept invariance) for both groups. Only two of the twenty four items on the PEAS did not meet strong invariance, and these items should be analyzed in future studies to see if the trends between parent groups remain. In light of the sample size and unequal distribution of groups, the good invariance and measurement characteristics of the revised PEAS are very positive. The invariance analysis provides further evidence that the PEAS is appropriate to use for both biological and nonbiological parent respondent groups.

Research Question 3

The third research question addressed the relationship of the PEAS subscales to overall parent satisfaction, as represented by the CSQ. It was expected that the three interpersonal subscales (Parent Assessor Relationship, Collaboration, and Child Assessor Relationship) would be most highly related to general parent satisfaction, and lower levels of Negative Feelings would be inversely related to general satisfaction.

This question was addressed both through correlation of the revised subscales with CSQ scores and Structural Equation Modeling showing the relationship between the various subscales and significant path loadings on general satisfaction. The correlation results showed that all of the revised PEAS subscales were significantly correlated with general satisfaction, and as hypothesized, the negative feelings subscale was negatively correlated with CSQ score. However, the subscale most strongly correlated with general

satisfaction was not the Parent Assessor Relationship and Collaboration subscale, but rather the New Understanding of Child (r = .64) subscale. The Child Assessor Relationship (r = .41) and Parent Assessor Relationship and Collaboration subscale (r = .48) had the next strongest correlations with CSQ score.

The structural equation model shows a more comprehensive depiction of how the PEAS subscales may interact and load on general satisfaction. The final model of the PEAS revision provided evidence that the PARC has direct effects on the Child Assessor Relationship, New Understanding of Child, and Negative Feelings subscales. Thus, rather than directly affecting parent satisfaction, PARC has indirect effects on general satisfaction through the other subscales. For example, if a direct path from PARC to CSQ score is drawn in the model, and the path from NUC to CSQ score is eliminated, the direct loading PARC on CSQ is significant. However, the model has a better fit with NUC having a path directly to CSQ, and eliminating the now non-significant PARC to CSQ path. Therefore, contrary to the hypothesis, it is a new understanding of the child that is most strongly associated with parent satisfaction. However it is important to note that the interpersonal relationship between the parent and assessor provides the foundation that than supplements the increased new understanding of the child.

The strong correlation between New Understanding of Child and parent satisfaction may help to fill in the gap between interpersonal variables, outcomes, and general satisfaction. As stated in the literature review, research has shown that outcomes have been less associated with general satisfaction than interpersonal variables. However, although parents want to be treated respectfully, positive interpersonal relationships are not the intended goal of the assessment: rather, parents have questions or concerns about

their child and are looking for answers. In Therapeutic Assessment, this new formulation or understanding of the child goes beyond a diagnosis and is referred to as the 'new story' for the child. As stated previously, the parent-assessor relationship and collaboration are seen as fundamental to then helping the parents come to a new understanding, both in terms of knowledge and compassion, for their child. If parents comes to a new understanding and have their questions answered, then they are more likely to be satisfied with an assessment.

Systemic Awarenss (SA) is the one subscale not significantly influenced by a direct path from PARC. Rather, is correlated with increases in Negative Feelings and New Understanding of Child. It is hypothesized that the positive correlation between Negative Feeling and SA is due to the increase in negative emotionality parents may experience when endorsing higher levels of family involvement in the child's problems.

The Systemic Awareness subscale is designed to reflect the parent's awareness of the larger systemic issues that may be contributing to or maintaining a child's difficulties. One of the goals of TA is to encourage parents to become a positive force in helping their child, and an increase in systemic awareness is seen as a desired step forward in TA. By increasing Systemic Awareness, parents then have more control and power in becoming 'unstuck' and helping their child. However, it appears that higher levels of Systemic Awareness may come with in increase in Negative Feelings. After reviewing the Systemic Awareness items, it is understandable that a parent endorsing increasing levels of 'causing' or culpability in their child's problems comes with an increase in negative feelings, such as guilt, feeling blamed, or like a bad parent.

Systemic Awareness is also positively correlated with increases in New Understanding of Child. Currently, these two subscales are correlated because it may be that a parent starts to understand their child differently, which then influences their view of the family, or the parent may become more aware of the family's role, which then leads to a different understanding of the child. In reality, these processes are most likely linked and iterative, in that they build upon each other throughout the assessment process.

The findings regarding the relationship between Systemic Awareness and the other subscales fit within the overall theoretical framework of TA. The holding environment created by the Parent-Assesor Relationhip and Collaboration is associated with decreased Negative Feelings and increased New Understanding of Child and perceptions of a positive Child Assessor Relationship.

Overall, the findings from the current study did not support the hypothesis that Parent Assessor Relationship and Collaboration would be most strongly associated with parent satisfaction. Rather, the parent-assessor relationship appears to indirectly affect satisfaction by serving as a foundational element which then increases a parent's New Understanding of Child. It is this new understanding that has the highest statistically significant direct effect on parent satisfaction. In addition, increases in Systemic Awareness are associated with an increase in both Negative Feelings and New Understanding of Child. Increasing Systemic Awareness may have a side effect of increasing Negative Feelings for parents. In contrast, increases in Systemic Awareness are also correlated with greater understanding of the child. By using structural equation modeling, significant direct effects among PEAS subscales demonstrate possible

interactions among these processes during a child psychological assessment and the resulting relationships with general satisfaction.

Research Question Four

The last research question for this study addresses the PEAS' ability to differentiate between collaborative and traditional assessment practices. Although the collaborative assessments had higher group means for the Parent Assessor Relationship, Child Assessor Relationship, New Understanding of Child, and Negative Feelings subscales than traditional assessment, the differences were not significant. It was hypothesized that the New Understanding of Child and Systemic Awareness subscales were the most likely to be significantly greater across assessment types, as these subscales reflect more unique goals of TA.

The Systemic Awareness was the only PEAS subscale found to be significantly higher for collaborative assessments. Increases in Systemic Awareness may not be a goal of more traditional assessment, but is considered part of the intervention aspect of Therapeutic Assessment. As discussed previously, it appears that increases in Systemic Awareness are associated with increases in Negative Feelings. The higher, but not yet significant, level of Negative Feelings for the collaborative assessments was unanticipated due to the reasoning that greater PARC should be associated with lower negative feelings. This primary reasoning and inverse relationship between negative feelings and the parent assessor relationship was demonstrated in the PEAS revision model; greater parent assessor relationship does have a direct effect on decreasing negative feelings. However, it appears that increases in Systemic Awareness and

corresponding increases in Negative Feelings is a corollary finding that should be addressed theoretically.

One explanation for the increase in Negative Feelings and its association with higher levels of Systemic Awareness is that parents who increasingly endorse family or systemic involvement then experience increased levels of guilt or blame. In fact, the increase in Negative Feelings associated with greater Systemic Awareness may be analogous to a similar pattern in increased parent reported child difficulties at the end of an assessment. A parent may initially underreport a child's difficulties at the beginning of an assessment or intervention, but then provide a more accurate, but technically worse, portrayal of the child by the end of the process. On the surface level, it appears as if the child's outcomes or symptoms have worsened over the course of treatment, however, this is often clinically interpreted as a correction from parents who may have underreported a child's difficulties at the beginning of the process and then gained a new awareness. The relationship between increased negative feelings and systemic awareness may be similar, in that greater awareness of the family's involvement may lead to increased negative feelings, which would indicate a less supportive experience. However, a strong parent-assessor relationship is also related to decreases of negative feelings, and so the dual influences of parent-assessor relationship and systemic awareness may work in tandem to keep negative feelings evoked by the assessment in check.

Overall, the predicted significant difference between traditional and collaborative assessment for Systemic Awareness was supported by the current findings. However, there were no significant differences between the two types of assessment for New Understanding of Child. It has been hypothesized that the relationships between Negative

Feelings should be interpreted within the context of increased Systemic Awareness and the inverse relationship with Parent Assessor Relationship and Collaboration within the theory of TA.

Implications for Therapeutic Assessment Theory and Satisfaction

PEAS Development

The Parent Experience of Assessment Scale (PEAS) has undergone a three year development and revision process. Based on the theory of Therapeutic Assessment, the PEAS began with qualitative statements from parents, expert sorting, and preliminary factor analysis resulting in six subscales and a total of 64 items. The pilot version was administered at 4 child assessment sites to provide variety in clients, settings, and assessment type. The revision has been made using Confirmatory Factor Analysis (CFA) and invariance (MG-MACS) testing to provide a high level of psychometric properties. The revised PEAS consists of a second order five subscale model with 24 items. Each subscale demonstrates adequate to good reliability, with Cronbach alphas ranging from .76 to .88.

In addition, the PEAS subscales had acceptable levels of skewness and kurtosis, thus helping prevent a 'ceiling effect' of responses. Items were developed specifically to curb response bias by making items which made it more difficult to select the extremes of Strongly Agree and Strongly Disagree (e.g. I learned a tremendous amount about my child from this assessment). Invariance testing found that the PEAS measures the same constructs and functions similarly across biological and nonbiological parent groups. Thus, the revised PEAS is a parent self report measure with good reliability and is

appropriate for use by clinicians and researchers interested in the parent experience of child assessment practice.

Relationship to Satisfaction

All five of the PEAS subscales are significantly correlated with general satisfaction, as measured by the most common satisfaction measure, the CSQ. Advantages of the PEAS include more nuanced understanding of different aspects of parents' experiences during their child's psychological assessment. The subscales provide information about both interpersonal experiences, through the Parent Assessor Relationship and Collaboration and Child Assessor Relationship subscales, and changes in parent knowledge or perspectives, through the New Understanding of Child and Systemic Awareness subscales. Structural equation modeling allowed a more in-depth look at the interactions among the PEAS subscales and direct and indirect effects on satisfaction. Rather than the parent assessor relationship directly affecting parent satisfaction, it appears to indirectly affect satisfaction through direct positive effects on the parent's perception of the Child Assessor Relationship and New Understanding of Child.

The high correlation between the New Understanding of Child and general satisfaction may help to fill in the gap in satisfaction research in terms of what may be most highly related to satisfaction. Previous research has found that outcomes, generally symptom reduction, and demographics are not significant predictors of satisfaction. In the current study, it was hypothesized that subscales representing interpersonal relationships would be the most highly related to parent satisfaction. However, it appears that parents endorsing higher levels of a new understanding of their child had the highest

direct effect on satisfaction. It may be that this new understanding, the new 'story' of the child in TA, may bridge the gap between interpersonal experiences, outcomes, and satisfaction. The items that were retained on the revised New Understanding of Child scale describe an enhanced understanding of the child, more realistic expectations, better communication, and new ideas for parenting. This shift in parent perspective may then lead to follow-through on recommendations and implementations of behavioral interventions. Thus, it may not be the outcomes themselves that lead to satisfaction, nor a strong relationship with the assessor in and of itself, rather, it is the new understanding gained from the assessment and a sense of being 'unstuck' which is most related to satisfaction. Simply put, providing answers to parents' assessment questions in a way that leads to new ideas and enhanced understanding of their child, not just providing a diagnosis or recommendations, are key components to predicting parent satisfaction with child assessment services.

In addition, satisfaction research has recently shifted to focusing on dissatisfaction, often defined as the lowest percentiles of respondents. Research has shown that unless negative experiences are specifically sought or targeted, respondents often fail to provide this information. The Negative Feelings subscale explicitly seeks to ascertain negative experiences parents may have regarding their child's assessment. Ranges for below average, average, and above average scores on the PEAS subscales have been provided based on the current sample. The PEAS subscales have demonstrated significant relationships to general satisfaction and may help inform and supplement program evaluation, training, and research related to consumer satisfaction of child assessment services.

Advancement of Therapeutic Assessment Theory

The PEAS subscales were developed based on Therapeutic Assessment theory, which highlights the importance of a strong client-assessor relationship and collaboration during the assessment process that then facilitate greater assimilation and depth of assessment findings for clients. Each subscale is designed to measure an underlying process hypothesized to lead to positive outcomes and change from a child TA. It was expected that the subscales would be correlated; however, the second order CFA model has gone a step further. The hierarchical model demonstrates there is a general factor measured by the PEAS subscales that accounts for their correlations. The finding that Parent Assessor Relationship and Collaboration (PARC) subscale serves as the second order factor provides statistical evidence that the parent assessor relationship truly is the fundamental, or overarching factor, measured by the PEAS and which affects the other processes/subscales.

The second order model and structural equation analyses also provide enhanced understanding of the interactions among the PEAS subscales and the processes they are designed to measure. As the hierarchical factor, PARC has significant direct effects on other processes in the assessment. Increases in PARC scores indicate an increase in Child-Assessor Relationship, New Understanding of Child, and a decrease in Negative Feelings towards the assessment. These results fit with the theory of TA, which suggests that a strong client assessor relationship is needed to create a safe holding environment to enhance the ability of the client to join the assessor on an 'observation deck' to explore and integrate deeper levels of assessment findings.

The finding that level of Systemic Awareness was significantly greater for collaborative assessment than traditional/best practice assessments fits with TA theory as well. Systemic Awareness is one of the intervention aspects of child TA, where parents are able to increasingly understand how their actions can help their child deal with his or her difficulties. For families that are 'stuck' and parents who have 'tried everything', finding that they can be part of the solution often engenders hope and recommitment of energy to the child. The greater level of Systemic Awareness is one of the unique aspects of TA and the significantly higher levels of Systemic Awareness scores from parents receiving collaborative assessment supports this claim.

Other findings from the PEAS revision particularly relevant to TA theory is the increase in negative feelings associated with increase in systemic awareness. Explanation of this finding is where theory and data come together in a unique way. For example, in the revised version of the PEAS scale, there is a significant correlation between Systemic Awareness and Negative Feelings. In the structural equation model using the revised PEAS, a choice was made to use a directional arrow from Systemic Awareness to Negative Feelings, rather than an arrow pointing the other direction or nondirectional correlation. One of the dangers of structural equation modeling is that model fit cannot tell the researcher if they have an arrow drawn in the wrong direction, and hence, solid theory is needed to guide these decisions.

In this case, the decision to draw the arrow from Systemic Awareness to Negative Feelings, indicating a direct effect from one to the other, was based on TA theory. One of the purposes of a strong assessor-client relationship is to create a safe environment to help contain negative or difficult feelings. Unconditional support and regard then allows

clients to lower defenses and truly 'hear' information that would normally be threatening to their worldview or sense of self. In the case of TA with children, it is not uncommon for parents to come to an assessment scapegoating a child – blaming the child's problems on inherent temperament or child characteristics. It is only once parents feel supported through the collaborative process where they start shifting their perspective of the child and, hopefully, see their role in the child's difficulties. This systemic aspect is often crystallized in a family intervention session, where the problem behavior is brought into the room and then systemic ways of dealing with the problem are explored. Hence, it was considered highly unlikely that an increase in negative feelings, making parents feel guilty or blamed, would lead to an increase in Systemic Awareness. In fact, just the opposite would be expected; if parents felt blamed for their child's problems early on, their defenses will be activated and they are then less likely to incorporate a more systemic view and more likely to maintain a view of the child as 'the problem'.

The three-way interplay between the Parent Assessor Relationship, Negative Feelings, and Systemic Awareness as measured by the PEAS seems to be highlighted in the relationships among these subscales. It appears there is a balancing act, with strong parent assessor relationship decreasing negative feelings to where systemic awareness can be explored. As systemic awareness increases, there is an increase in negative feelings as well. Based on TA theory, it is probable that the strong parent-assessor relationship is helping the parents to manage and contain those negative feelings. However, this may lead to a higher level of negative feelings for collaborative assessment that is initially counterintuitive, but seems to fit when taken in the context of significant increases in systemic awareness.

Lastly, the significant correlation between New Understanding of Child and Systemic Awareness reflects the common pattern in TA that parents shift in both their perceptions of the 'story' for their child (e.g., from 'bad' to 'sad') with a corresponding increase in understanding how the family plays a role in the child's problems, and maybe more importantly, how the family can intervene in positive directions. In the SEM model (Figure 3) this relationship was left as a correlation, rather than turned into a directional arrow that would imply one precedes or directly affects the other. It is conceivable that as parents begin to have a better understanding of their child they begin to have a more systemic perspective, but the reverse is also possible, that parents could begin to perceive the family role in a child's problems which then augments their new understanding of the child. Because the precedence of one of these processes has not been established, either empirically or theoretically, they remain correlated and are open to further investigation and revision.

Overall, the current study provides the development of the PEAS and the resulting good psychometric properties. In addition, it has contributed new evidence of direct and indirect effects of the PEAS subscales related to general satisfaction and the processes underlying Therapeutic Assessment.

Directions for Future Research

Further PEAS Versions

The current version of the PEAS is a post only measure, meaning that the wording is framed to assess parent reactions and experiences at the conclusion of an assessment. However, to further understand the underlying development of these processes, the verb tense of the PEAS could be changed to allow for repeated measurement. For example,

the PEAS could be administered after the assessment sessions of a Therapeutic Assessment and again after the feedback and/or follow up sessions. This could allow for greater understanding of the processes behind the subscales, such as whether shifts are seen first in New Understanding of Child or Systemic Awareness to further elucidate which precedes, and therefore influences, the other.

Invariance testing for different groups of respondents is also needed, particularly for any language translations. Items may not tap the same construct or be interpreted the same way across different cultures. Therefore, before comparisons between mean scores on different versions are made, invariance testing should verify the measurement aspects of other language versions.

Although greater emphasis and research is being placed on receiving parent feedback about child mental health services, the perspectives of children and adolescents are still underrepresented. The PEAS subscales and revision process could be used as a model to create child and adolescent versions. Many of the subscales could be modified to measure a youth's perspective of the assessment, such as the child assessor relationship, a level of new understanding of self, how well the child thinks the parents understand the child after the assessment, and negative feelings about the assessment. The child and adolescent experience of an assessment is an important perspective that can lay the foundation for expectations with future mental health services and should be an increasing priority for both clinicians and researchers.

Implications for Child Assessment Research

Although general satisfaction was used to help provide evidence of convergent construct validity, further validity studies for the PEAS constructs measured by the

subscales are needed. General satisfaction is a first-order or immediate outcome of an assessment, however, follow up studies are needed to measure second order outcomes, such as follow through on assessment recommendations and maintenance of outcomes. Other possible constructs related to PEAS subscales include alliance and family process measures. Alliance measures have been geared towards therapy, and hence were not worded in ways consistent with assessment. However, a study showing convergent correlations between an alliance measure and Parent Assessor Relationship and Collaboration would provide evidence of the similarity of the two constructs. Similarly, changes in family measures, such as family conflict and cohesion may be associated with changes in the Systemic Awareness and New Understanding of Child Subscales.

It may also be that different profiles of responding to assessments could be gleaned through differences in PEAS subscale score profiles. For example, some parents may report higher PEAS subscale scores at a three to six month follow-up session once the assessment process and results have had time to be practiced and worked into daily routine. Other parents may report higher subscale scores at the end of an assessment and then either maintain or decrease their scores at follow up.

Other questions include whether the correlation between the increase in negative feelings and increase in systemic awareness is a pattern consistent across collaborative assessments. It may be that in comprehensive child TA assessments where parents process the assessment sessions with an assessor, parents may get enough support to keep negative feelings lower for collaborative assessments. Similarly, the collaboration subscale was merged with the parent assessor relationship subscale in the revised PEAS. It may be that the collaborative items relating to more comprehensive Therapeutic

Assessments did not fit with the processes at the sites in the study. Revised items for collaboration could be included with the revised PEAS to continue investigating the collaborative piece of Therapeutic Assessment.

Limitations

Limitations related to data collection and sampling for the current study include missing CSQ data for one site in addition to uneven group sizes for the invariance testing between parent respondent groups. Although the sample size was above the minimum 100 cases recommended for confirmatory factor analysis, larger sample sizes would have provided even stronger data in terms of fit of models.

One of the strengths of the study was the variety of sites and clientele administered the pilot version of the PEAS. However, direct comparisons between sites or other demographics were not as feasible, as the variability confounded other questions. For example, any differences between sites on the PEAS subscales could be related to covariates such as SES, collection method of the PEAS, type of assessment, or respondent characteristics. Specific questions regarding these covariates would need to be explored with more homogenous or controlled samples. Specifically, severity of child disorder is one area that has been found to predict parent satisfaction and is an important area to investigate. Unfortunately, no severity ratings, such as GAF scores, were available nor were child diagnoses provided consistently. All four of the sites provide outpatient assessment services, and so assessments with inpatient populations may also be needed for further comparison.

Although variety in assessment settings, type of assessments, and parent respondents were sought to provide the variability needed for the PEAS, the self selection

of sites may have prevented more differences between standard practice/traditional and collaborative/therapeutic assessment from being found. Because all four of the sites were willing to supplement their practice or were already practicing aspects of therapeutic assessment, this may indicate that their level of 'standard practice' is already more collaborative in nature than typical assessment services.

Summary

The Parent Experience of Assessment Scale (PEAS) was created based on the theoretical orientation of Therapeutic Assessment to measure the processes hypothesized to contribute to successful Therapeutic Assessments. These processes include the interpersonal aspects of a strong parent-assessor relationship, collaboration with the parents throughout the assessment, and a positive child-assessor relationship. The support provided by these processes then facilitates shifts in the parents' understanding of their child and greater awareness of family or systemic roles in the child's difficulties. Negative feelings evoked by the assessment process are to be minimized through the supportive interpersonal relationship with the assessor.

Current research in Therapeutic Assessment has focused on repeated measures targeting child symptomology, family processes, and positive/negative emotionality to provide evidence of the effectiveness of TA-C. The most recent research has been focusing on disseminating components of child TA into real world practice and evaluating whether it significantly enhances the assessment process. Although previous work has been conducted in determining adult experiences with assessment, no quantifiable measure of the constructs underlying the TA-C process had been developed. Hence, the development of the PEAS provides an important link in establishing

quantifiable constructs that can reliably measure parent experiences during a child psychological assessment.

The satisfaction literature has increasingly focused on parent satisfaction with child mental health services (e.g. therapy and intervention), but has not specifically targeted child assessment services. In addition, a long history of criticisms have followed general satisfaction research including a lack of cohesive theory, poor psychometrics and scale development, sampling issues, the ceiling effect of client responses, and the lack of strong predictors for parent satisfaction. Research on satisfaction has increasingly found the interpersonal interactions during child services may be more related to satisfaction than outcomes or demographics.

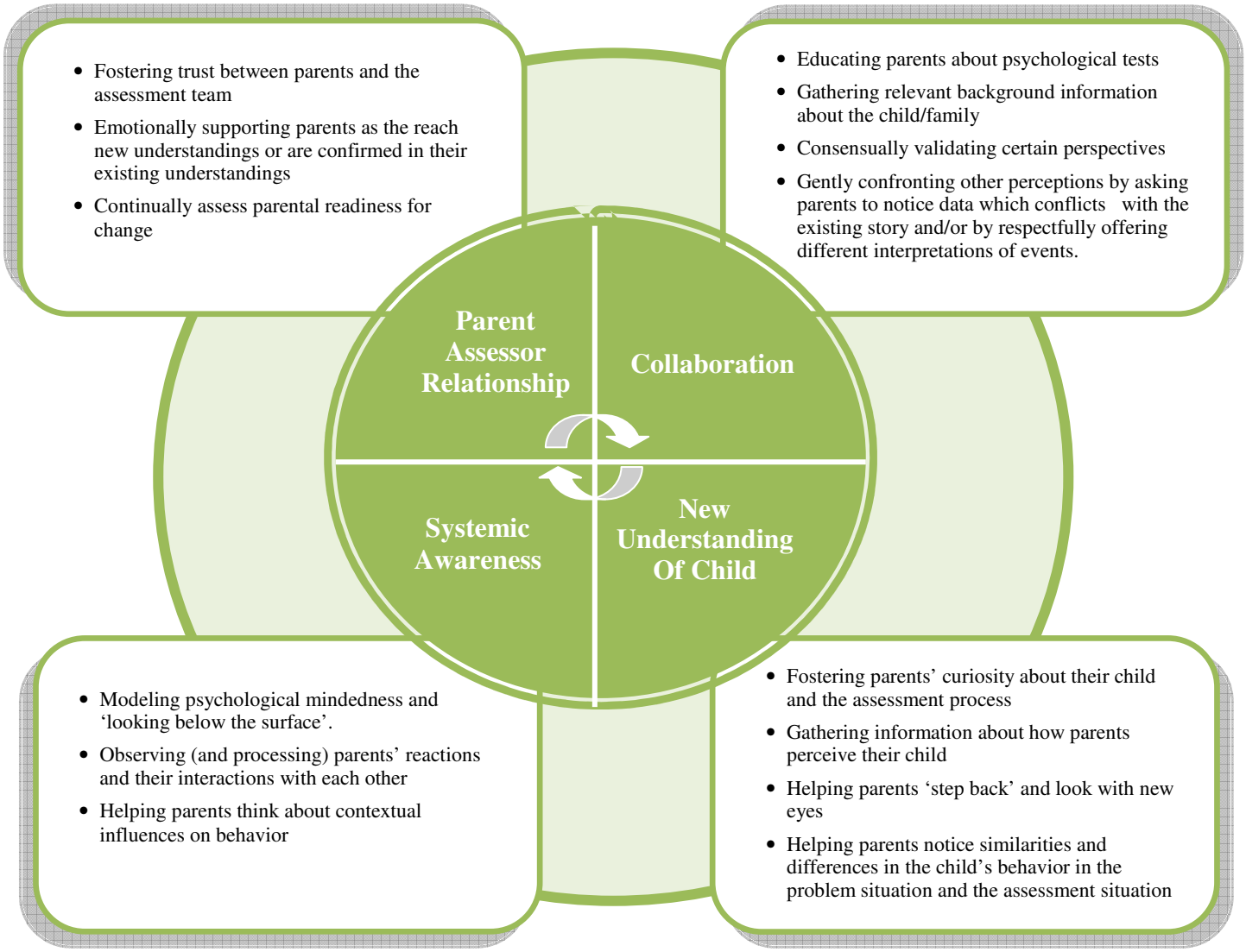
The findings of the current study have the potential to advance both the areas of Therapeutic Assessment and parent satisfaction research. The PEAS consists of five subscales and was revised via a more psychometrically sound process of confirmatory factor analysis and invariance testing across parent respondent groups. Furthermore, the current study has provided evidence supporting the theoretical hypotheses of TA, such as demonstrating that the Parent Assessor Relationship and Collaboration subscale is a hierarchical factor that directly affects three of the other PEAS subscales. In addition, the study found that it is the New Understanding of Child, facilitated by a strong parent assessor relationship, which has the most significant, direct effect on parent satisfaction.

Overall, the revised PEAS is a parent self report measure with good psychometric properties and a demonstrated relationship to parent satisfaction. The PEAS has been designed for use in measuring parents' experiences in both traditional and collaborative assessments, with various parent respondent groups, and in diverse assessment settings. It

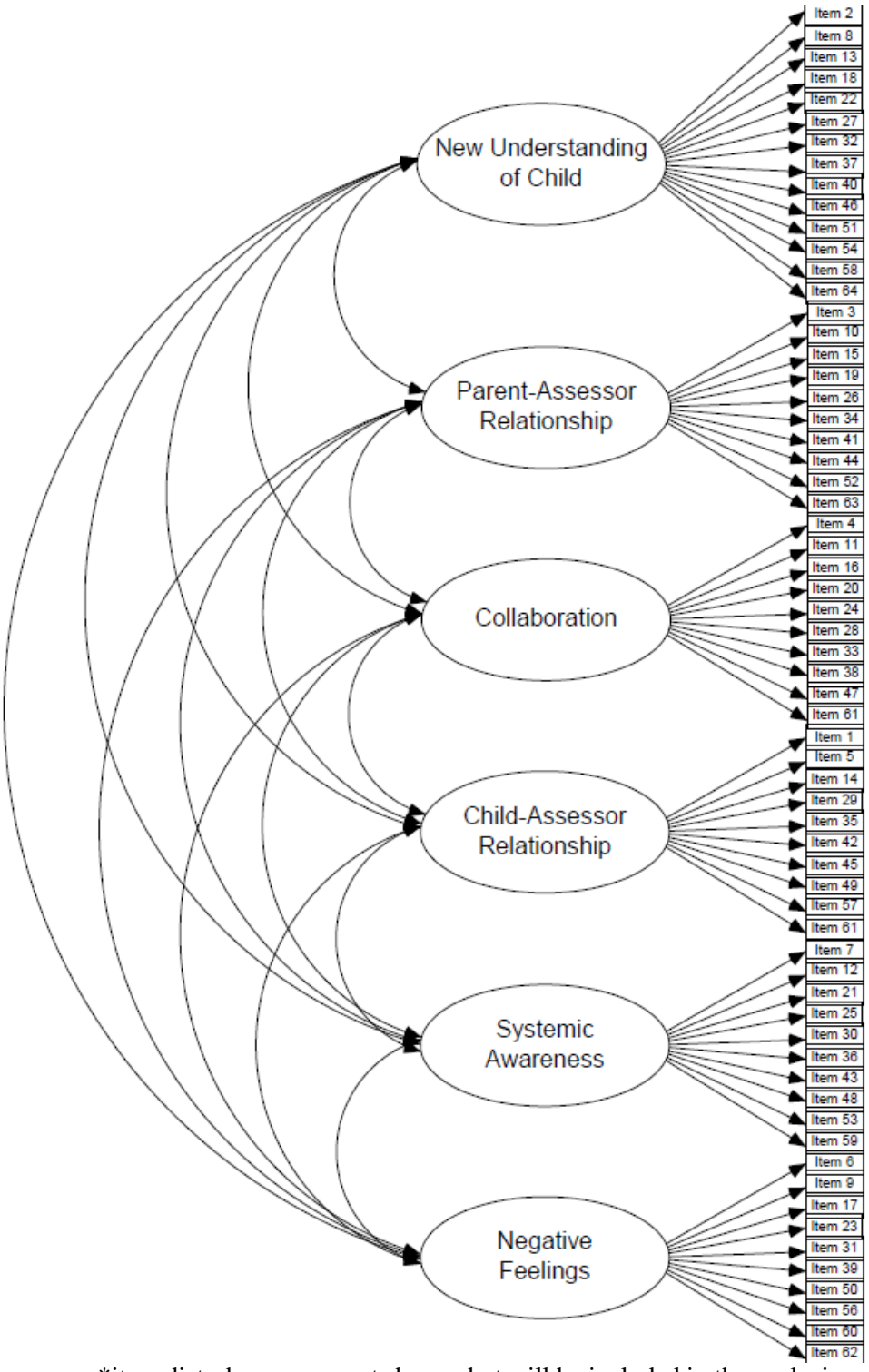
can provide more specific feedback about how parents experience their child's psychological assessment process for child assessment practitioners and researchers alike. Thus, the PEAS not only helps quantify the underlying constructs hypothesized to take place in collaborative/therapeutic assessment, but has also provided insight into the interactions among these processes, setting the stage for further investigation.

Appendices

Appendix A
Conceptual Diagram of Parent Interpersonal Processes in TA-C

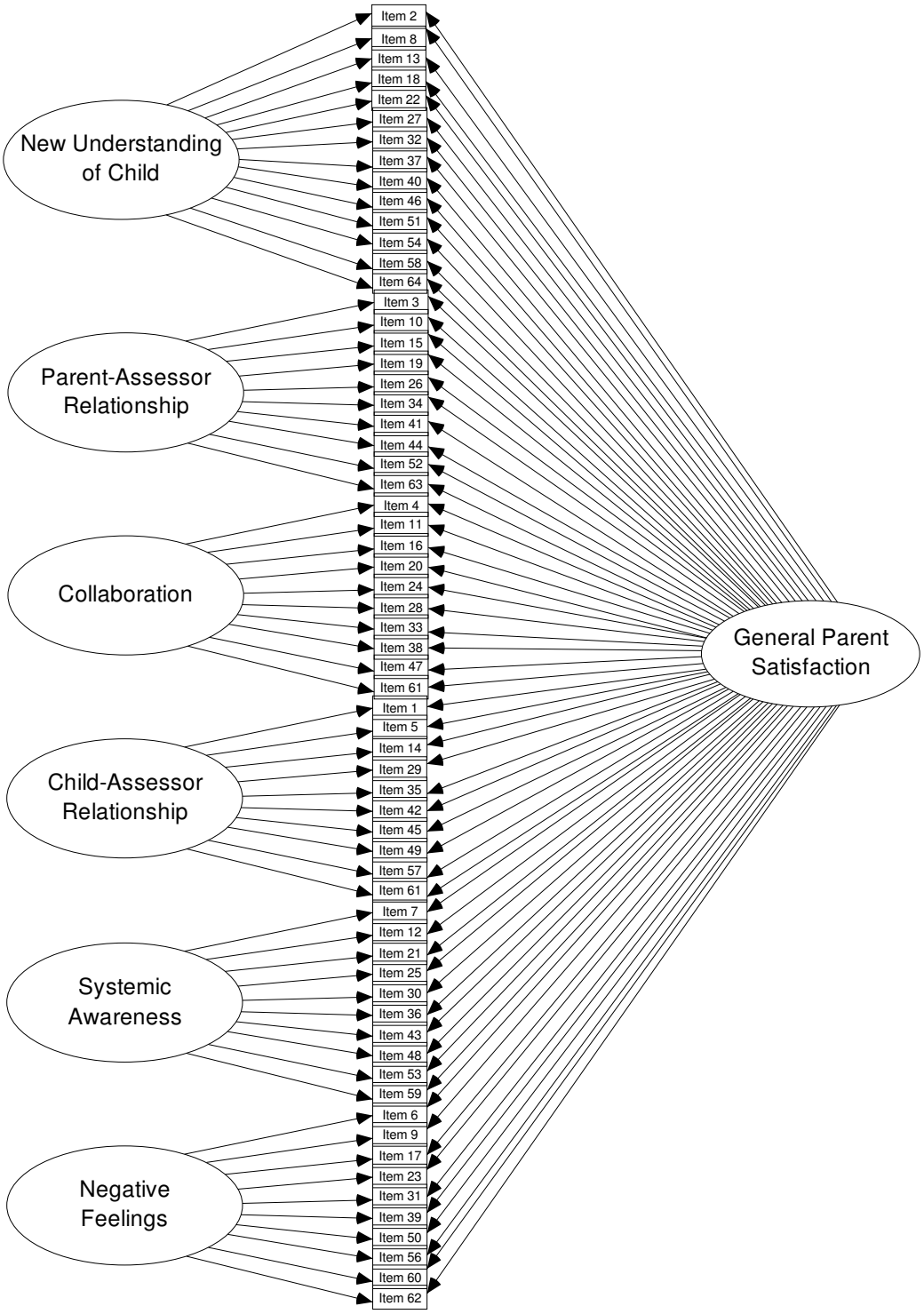


Appendix B
First Order Six Factor Model of PEAS



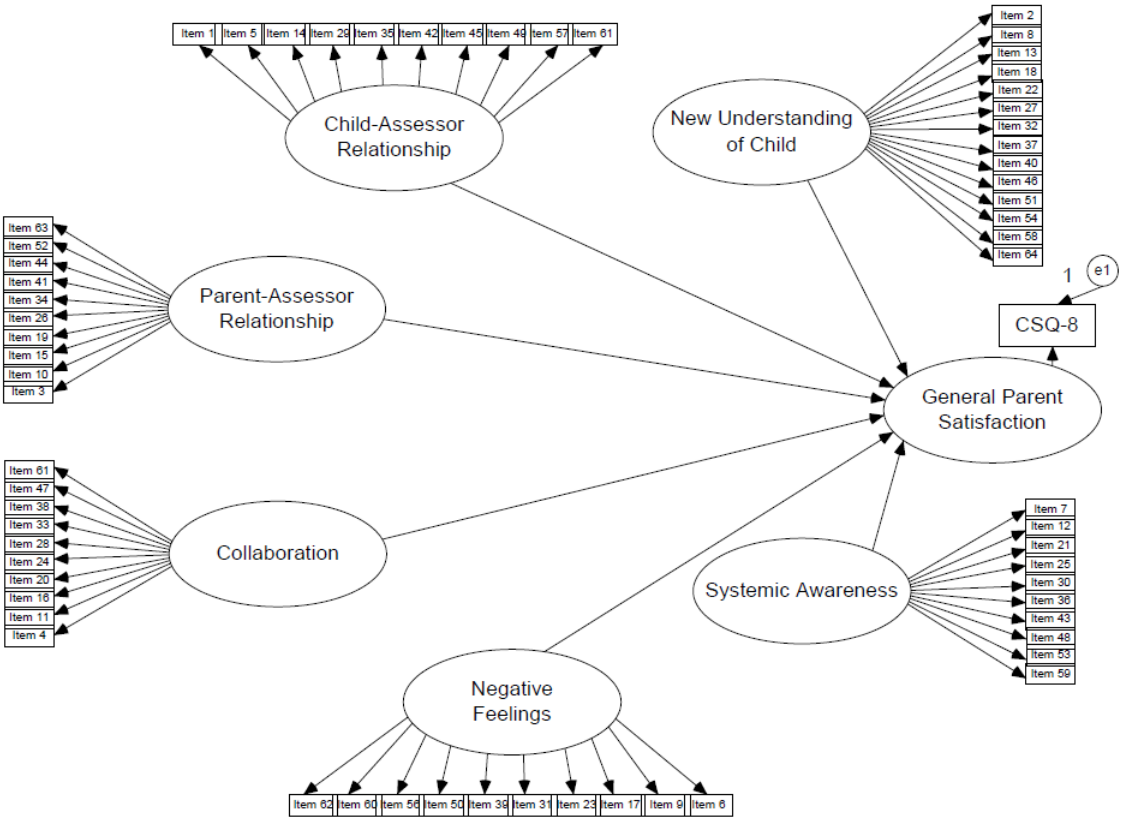
*item disturbances are not shown, but will be included in the analysis

Appendix C
Nested (Bi-factor) Model Confirmatory Factory Analysis for PEAS



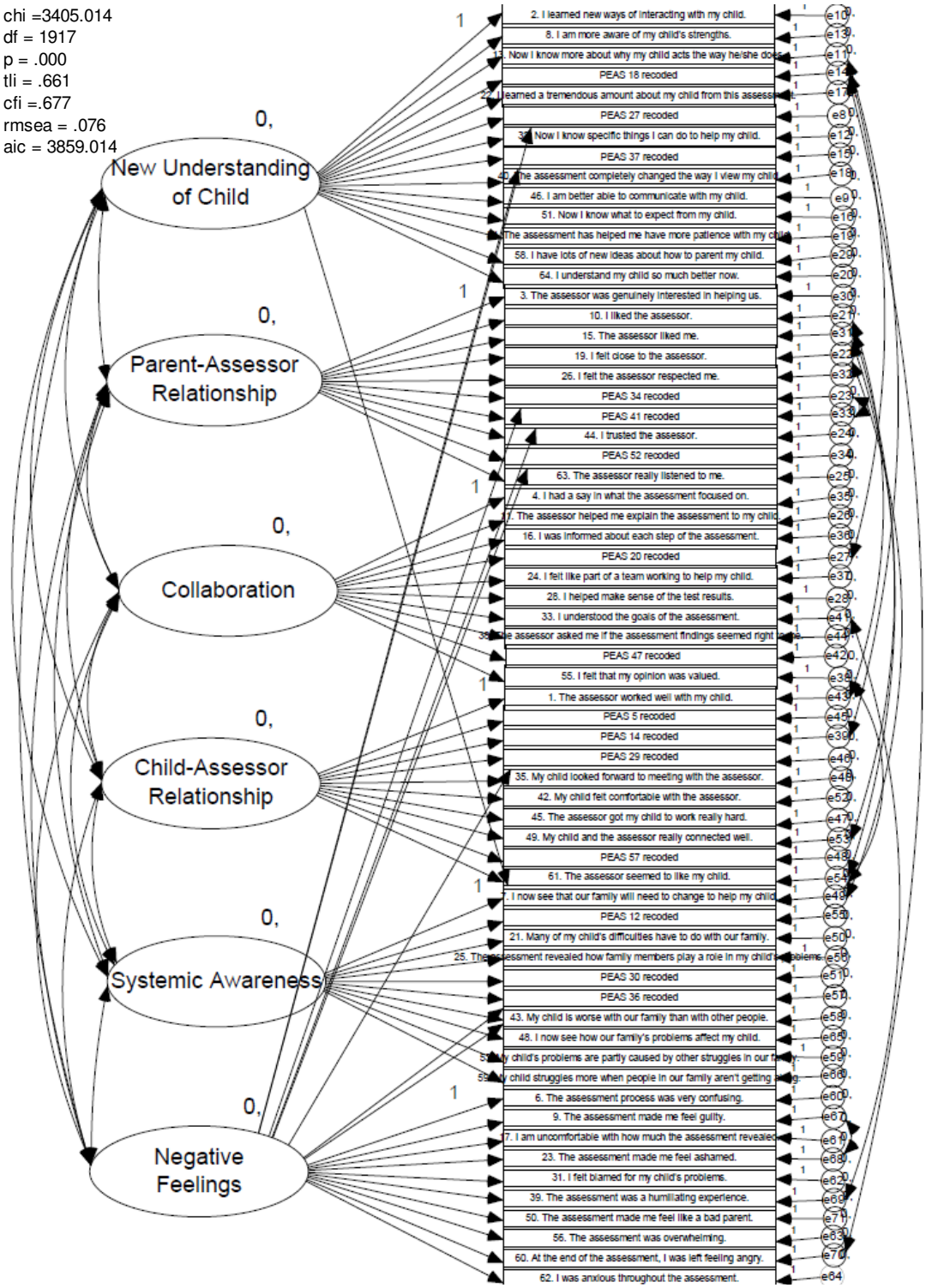
Appendix D

Structural Equation Model of Direct Effects of PEAS Subscales on General Satisfaction

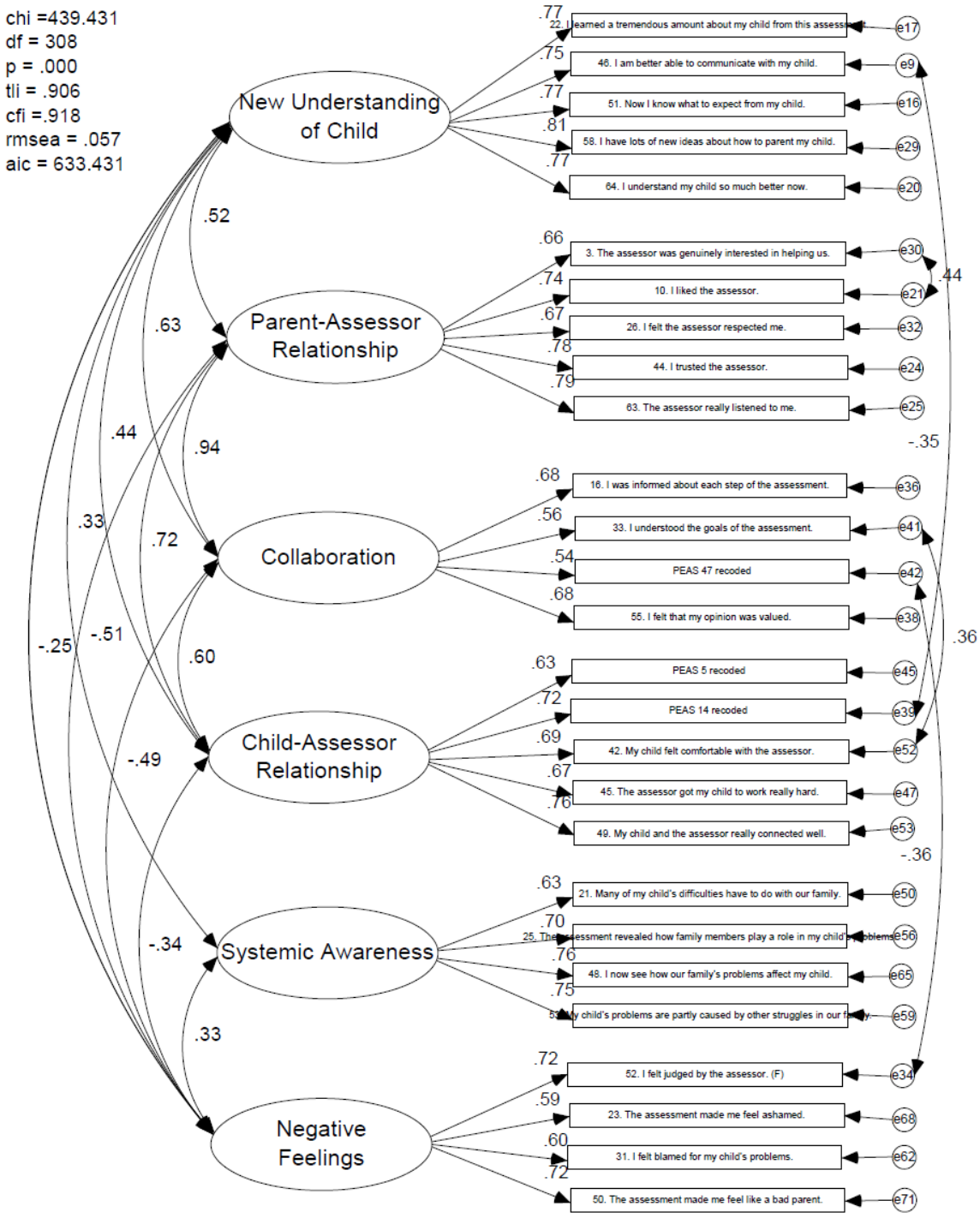


*item disturbances are not shown, but will be included in the analysis

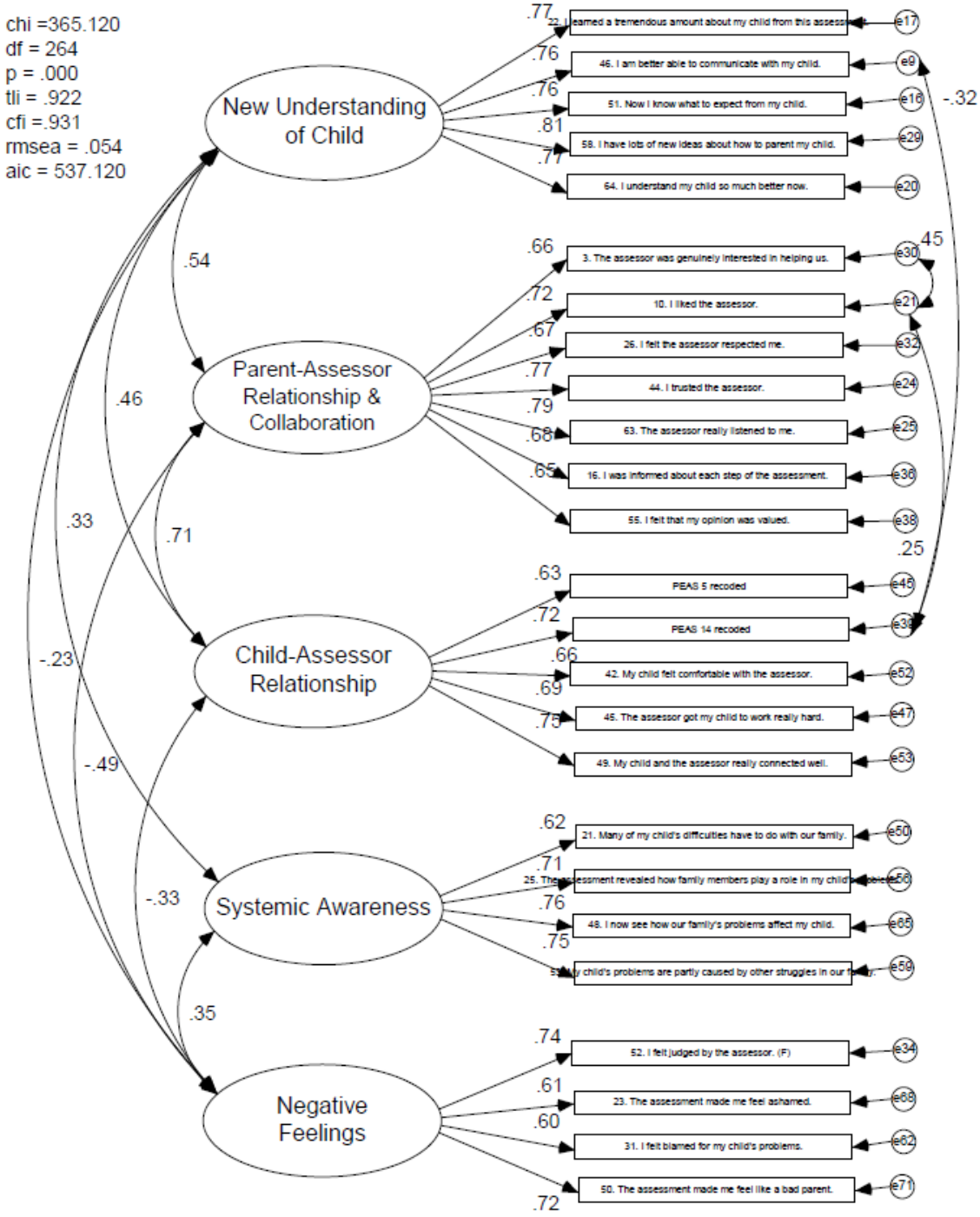
Appendix E
First Order Base Model



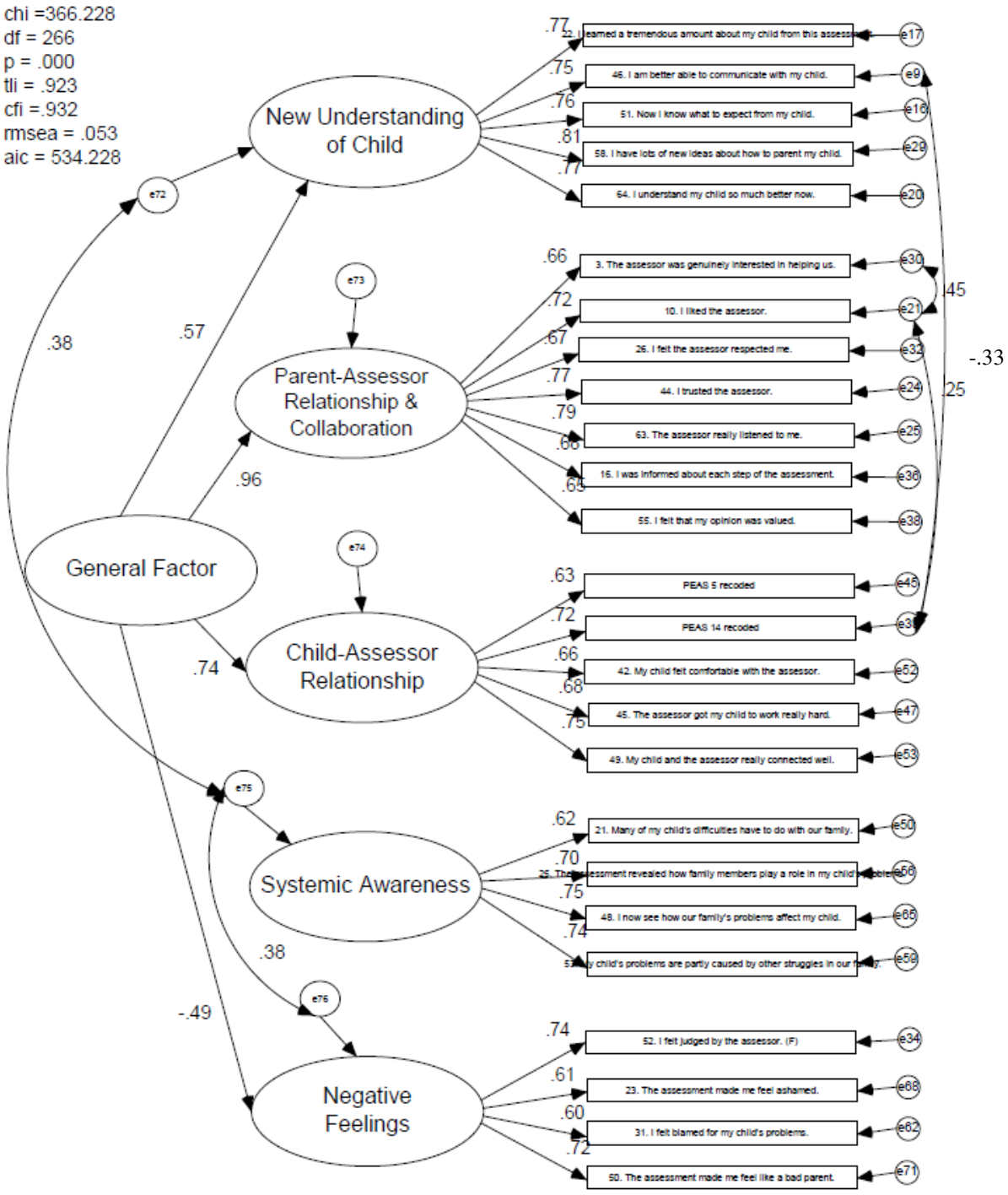
Appendix F
First Order 27 Item Model



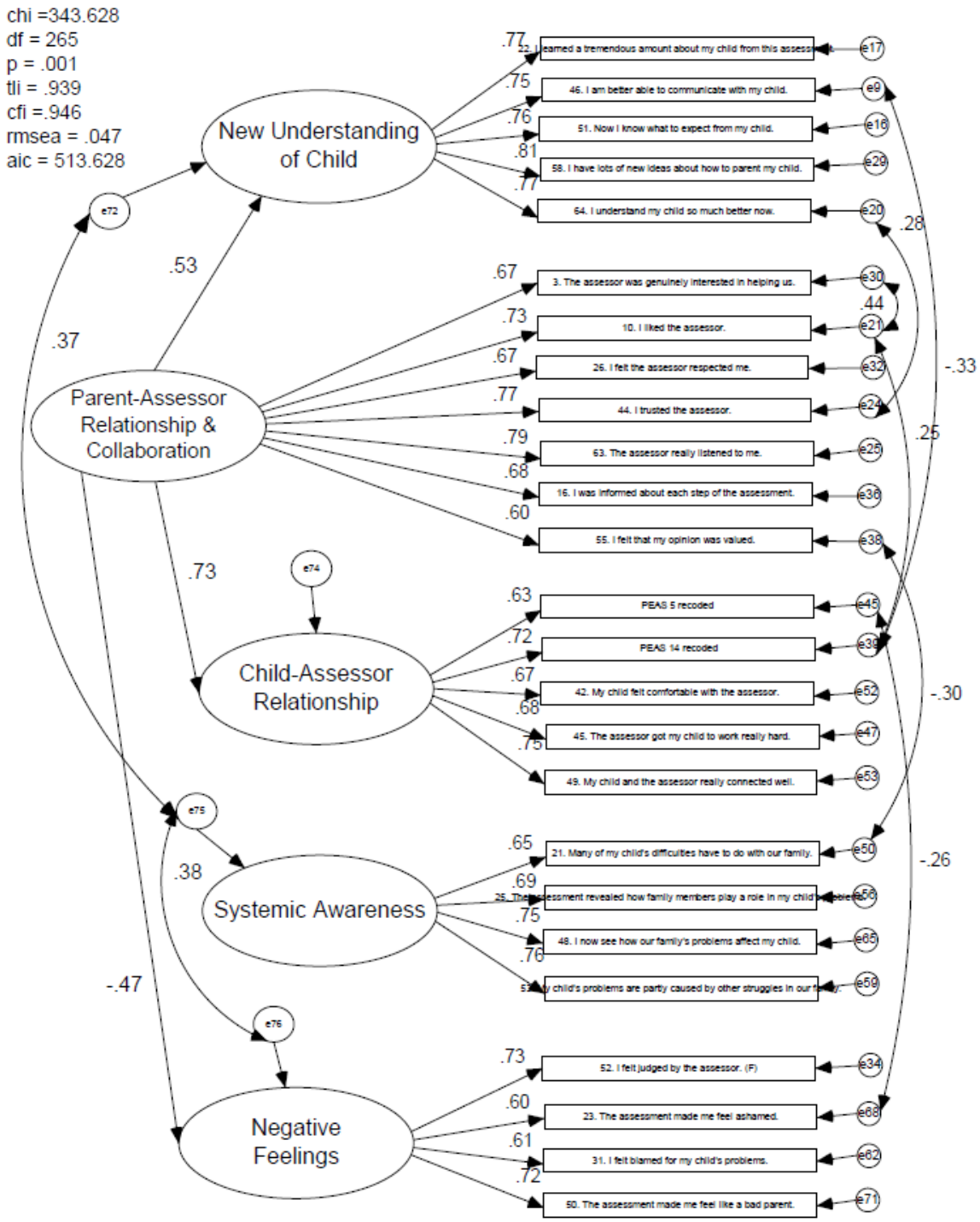
Appendix G
First Order Five Factor Model (25 Items)



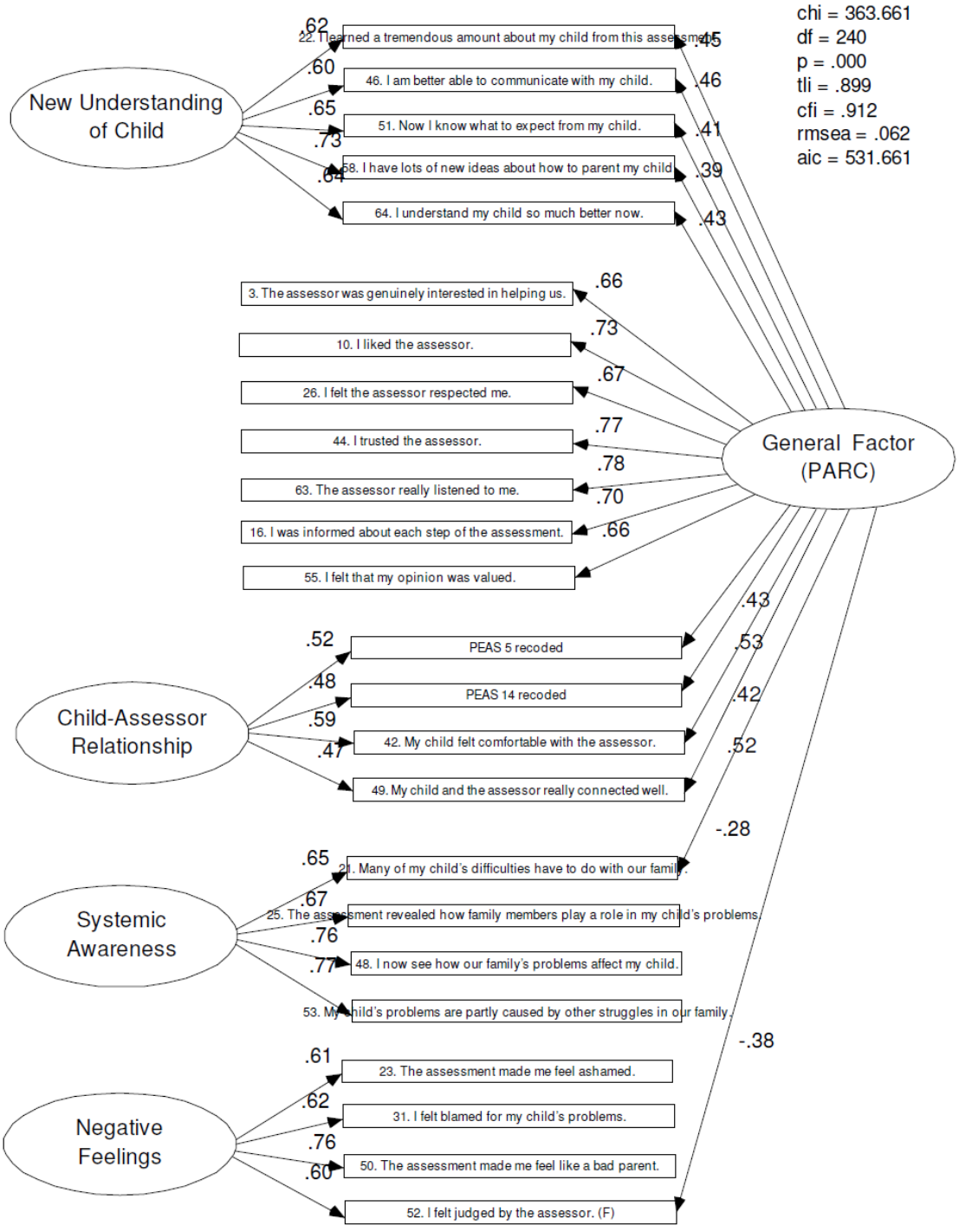
Appendix H
Second Order Model with General Factor



Appendix I
Second Order PARC Best Model



Appendix J
Bi-Factor Model



Appendix K
Demographic Analysis Results

Table 22: Analysis of Variance for Child Ethnicity

Scale	Sum of Squares	*df*	*MS*	*F*	*p*
CSQ Score					
Between Groups	.807	3	.269	1.079	.362
Within Groups	21.690	87	.249		
Total	22.497	90			
New Understanding of Child					
Between Groups	.361	3	.120	.181	.909
Within Groups	69.705	105	.664		
Total	70.066	108			
Parent Assessor Relationship and Collaboration					
Between Groups	1.168	3	.389	1.303	.277
Within Groups	31.361	105	.299		
Total	32.529	108			
Child Assessor Relationship					
Between Groups	2.168	3	.723	1.505	.218
Within Groups	50.430	105	.480		
Total	52.599	108			
Systemic Awareness					
Between Groups	1.210	3	.403	.472	.702
Within Groups	89.644	105	.854		
Total	90.854	108			
Negative Feelings					
Between Groups	2.776	3	.925	2.116	.103
Within Groups	45.913	105	.437		
Total	48.688	108			

Table 23: Independent t-tests for Gender

Respondent Gender

Test/Subscale	Female		Male		*t*	*df*	*p*
	M	**SD**	**M**	**SD**			
CSQ	3.63	.532	3.67	.369	-2.75	113	.784
Parent Assessor Relationship and Collaboration	4.36	.553	4.41	.421	-.447	132	.656
New Understanding of Child	3.71	.812	3.88	.596	-1.038	132	.301
Child Assessor Relationship	4.19	.727	4.41	.523	.167	132	.867
Systemic Awareness	2.78	.952	2.62	.884	.807	132	.421
Negative Feelings	1.64	.651	1.59	.640	.341	132	.734

Child Gender

Test/Subscale	Female		Male		*t*	*df*	*p*
	M	**SD**	**M**	**SD**			
CSQ	3.61	.541	3.67	.479	-.552	89	.582
Parent Assessor Relationship and Collaboration	4.35	.537	4.36	.571	.138	103	.890
New Understanding of Child	3.74	.816	3.74	.809	-.007	103	.994
Child Assessor Relationship	4.24	.740	4.08	.680	1.085	103	.280
Systemic Awareness	2.74	.810	2.90	.989	-.830	103	.408
Negative Feelings	1.69	.748	1.71	.629	-.105	103	.916

Appendix L
Reliability Statistics for Revised PEAS

Table 24: Reliability Statistic for Parent Assessor Relationship and Collaboration

Items	Item-Total Statistics			
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 3	25.96	10.953	.669	.862
Item 10	26.11	10.315	.720	.854
Item 16	26.48	9.259	.646	.872
Item 26	26.26	10.803	.629	.866
Item 44	26.12	10.626	.721	.856
Item 55	26.18	10.654	.615	.867
Item 63	26.20	10.068	.726	.853

Scale Statistics				
Mean	Variance	SD	N of Items	Cronbach's Alpha
30.55	13.852	3.722	7	.879

Table 25: Reliability Statistics for New Understanding of Child

Items	Item-Total Statistics			
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 22	14.82	9.727	.708	.861
Item 46	15.04	10.284	.693	.864
Item 51	14.86	10.077	.715	.859
Item 58	15.25	9.574	.761	.848
Item 64	14.89	9.694	.721	.858

Scale Statistics				
Mean	Variance	SD	N of Items	Cronbach's Alpha
18.72	14.987	3.871	5	.883

Table 26: Reliability Statistics for Child Assessor Relationship

Items	Item-Total Statistics			
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 5	12.34	4.633	.589	.748
Item 14	12.50	4.297	.621	.733
Item 42	12.54	4.791	.614	.738
Item 49	12.85	4.534	.591	.747

Scale Statistics				
Mean	Variance	SD	N of Items	Cronbach's Alpha
16.75	7.589	2.755	4	.793

Table 27: Reliability Statistics for Systemic Awareness

Items	Item-Total Statistics			
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 21	8.69	8.942	.554	.770
Item 25	7.95	8.596	.574	.761
Item 48	7.98	8.439	.650	.725
Item 53	8.51	7.479	.658	.720

Scale Statistics				
Mean	Variance	SD	N of Items	Cronbach's Alpha
11.05	13.922	3.731	4	.796

Table 28: Reliability Statistics for Negative Feelings

Items	Item-Total Statistics			
	Scale Mean if Item Deleted	**Scale Variance if Item Deleted**	**Corrected Item-Total Correlation**	**Cronbach's Alpha if Item Deleted**
Item 23	4.81	3.927	.529	.715
Item 31	4.84	3.938	.527	.715
Item 50	4.88	3.986	.612	.667
Item 52	5.03	4.420	.561	.700

Scale Statistics				
Mean	**Variance**	**SD**	**N of Items**	**Cronbach's Alpha**
6.52	6.687	2.586	4	.756

Appendix M
PEAS Total Score Conversion to T-Scores

Table 29: Subscale Total Score Conversion to T-Scores

T-Score	Subscale Total Score				
	Parent Assessor Relationship & Collaboration	**New Understanding of Child**	**Child Assessor Relationship**	**Systemic Awareness**	**Negative Feelings**
≥ 75					≥ 13
74				20	
73					
72					
71				19	12
70					
69				18	
68					11
67					
66		25		17	
65					
64		24			10
63				16	
62	35		20		
61		23		15	
60					9
59	34				
58		22	19	14	
57	33				
56		21			8
55			18	13	
54	32				
53		20		12	
52					7
51	31	19	17		
50				11	

49					
48	30	18			
47			16	10	6
46	29	17			
45				9	
44			15		5
43	28	16			
42				8	
41					
40	27	15	14		4
39				7	
38	26	14			
37				6	
36			13		
35	25	13			
34				5	
33		12	12		
32	24				
31				4	
30	23	11			
29			11		
28					
27	22	10			
26					
25		9	10		
24	21				
23					
22		8	9		
21					
≤ 20	≤ 21	≤ 7	≤ 8		

Appendix N
Client Satisfaction Questionnaire (CSQ-8)

Please help us improve our program by answering some questions about the assessment your child received. We are interested in your honest opinion, whether it is positive or negative. Please answer all of the questions. Thank you very much, we really appreciate your help.

CIRCLE YOUR ANSWER

1. How would you rate the quality of service you received?

1	*2*	*3*	*4*
Poor	*Fair*	*Good*	*Excellent*

2. Did you get the kind of service you wanted?

1	*2*	*3*	*4*
Yes, definitely	*Yes, generally*	*No, not really*	*No, definitely not*

3. To what extent has our program met your needs?

1	*2*	*3*	*4*
None of my needs have been met	*Only a few of my needs have been met*	*Most of my needs have been met*	*Almost all of my needs have been met*

4. If a friend were in need of similar help, would you recommend our program to him or her?

1	*2*	*3*	*4*
Yes, definitely	*Yes, generally*	*No, not really*	*No, definitely not*

5. How satisfied are you with the amount of help you have received?

1	*2*	*3*	*4*
Very satisfied	*Mostly satisfied*	*Indifferent or mildly dissatisfied*	*Quite dissatisfied*

6. Have the services you received helped you to deal more effectively with your child's problems?

1	*2*	*3*	*4*
No, they seemed to make things worse	*No, they really didn't help*	*Yes, they helped somewhat*	*Yes, they helped a great deal*

7. In an overall, general sense, how satisfied are you with the service you have received?

1	*2*	*3*	*4*
Quite dissatisfied	*Indifferent of mildly dissatisfied*	*Mostly satisfied*	*Very satisfied*

8. If you were to seek help again, would you come back to our program?

1	*2*	*3*	*4*
Yes, definitely	*Yes, generally*	*No, not really*	*No, definitely not*

Appendix O

PARENT EXPERIENCE OF ASSESSMENT SCALE (PEAS)

Respondent's Name	
Child's Name	
Date	

This questionnaire deals with your thoughts and feelings about your child's psychological assessment. Please read each statement carefully. Once you decide how much you agree or disagree with a statement, circle the number that best matches how the statement applies to you.

If you believe an items does not apply to you, please mark 'Neutral' category for that item. Be as honest and as accurate as possible. Please do not skip any item and check only one box for each statement.

Use the following scale to rate each statement:	**Strongly Disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
1. The assessor worked well with my child.	1	2	3	4	5
2. I learned new ways of interacting with my child.	1	2	3	4	5
3. The assessor was genuinely interested in helping us.	1	2	3	4	5
4. I had a say in what the assessment focused on.	1	2	3	4	5
5. My child did not like the assessor.	1	2	3	4	5
6. The assessment process was very confusing.	1	2	3	4	5
7. I now see that our family will need to change to help my child.	1	2	3	4	5
8. I am more aware of my child's strengths.	1	2	3	4	5
9. The assessment made me feel guilty.	1	2	3	4	5
10. I liked the assessor.	1	2	3	4	5

11. The assessor helped me explain the assessment to my child.	1	2	3	4	5
12. Our family has little to do with why my child has problems.	1	2	3	4	5
13. Now I know more about why my child acts the way he/she does.	1	2	3	4	5
14. My child never really warmed up to the assessor.	1	2	3	4	5
15. The assessor liked me.	1	2	3	4	5
16. I was informed about each step of the assessment.	1	2	3	4	5
17. I am uncomfortable with how much the assessment revealed.	1	2	3	4	5
18. I didn't learn anything new about my child from the assessment.	1	2	3	4	5
19. I felt close to the assessor.	1	2	3	4	5
20. I never really understood the point of the assessment.	1	2	3	4	5
21. Many of my child's difficulties have to do with our family.	1	2	3	4	5
22. I learned a tremendous amount about my child from this assessment.	1	2	3	4	5
23. The assessment made me feel ashamed.	1	2	3	4	5
24. I felt like part of a team working to help my child.	1	2	3	4	5
25. The assessment revealed how family members play a role in my child's problems.	1	2	3	4	5
26. I felt the assessor respected me.	1	2	3	4	5
27. Now I am more confused about how to handle my child.	1	2	3	4	5
28. I helped make sense of the test results.	1	2	3	4	5
29. The assessor never really	1	2	3	4	5

understood my child.					
30. I don't believe our family makes my child's problems worse.	1	2	3	4	5
31. I felt blamed for my child's problems.	1	2	3	4	5
32. Now I know specific things I can do to help my child.	1	2	3	4	5
33. I understood the goals of the assessment.	1	2	3	4	5
34. I felt the assessor was cold towards me.	1	2	3	4	5
35. My child looked forward to meeting with the assessor.	1	2	3	4	5
36. My child is the only person in our family who needs to change.	1	2	3	4	5
37. I wish I had learned more concrete ways to help my child day to day.	1	2	3	4	5
38. The assessor asked me if the assessment findings seemed right to me.	1	2	3	4	5
39. The assessment was a humiliating experience.	1	2	3	4	5
40. The assessment completely changed the way I view my child.	1	2	3	4	5
41. I felt the assessor looked down on me.	1	2	3	4	5
42. My child felt comfortable with the assessor.	1	2	3	4	5
43. My child is worse with our family than with other people.	1	2	3	4	5
44. I trusted the assessor.	1	2	3	4	5
45. The assessor got my child to work really hard.	1	2	3	4	5
46. I am better able to communicate with my child.	1	2	3	4	5
47. No one ever told me what would happen during the	1	2	3	4	5

assessment.					
48. I now see how our family's problems affect my child.	1	2	3	4	5
49. My child and the assessor really connected well.	1	2	3	4	5
50. The assessment made me feel like a bad parent.	1	2	3	4	5
51. Now I know what to expect from my child.	1	2	3	4	5
52. I felt judged by the assessor.	1	2	3	4	5
53. My child's problems are partly caused by other struggles in our family.	1	2	3	4	5
54. The assessment has helped me have more patience with my child.	1	2	3	4	5
55. I felt that my opinion was valued.	1	2	3	4	5
56. The assessment was overwhelming.	1	2	3	4	5
57. My child dreaded almost every meeting with the assessor.	1	2	3	4	5
58. I have lots of new ideas about how to parent my child.	1	2	3	4	5
59. My child struggles more when people in our family aren't getting along.	1	2	3	4	5
60. At the end of the assessment, I was left feeling angry.	1	2	3	4	5
61. The assessor seemed to like my child.	1	2	3	4	5
62. I was anxious throughout the assessment.	1	2	3	4	5
63. The assessor really listened to me.	1	2	3	4	5
64. I understand my child so much better now.	1	2	3	4	5

Appendix P
Final Version of the PEAS

The Parent Experience of Assessment Scale					
Statement	**Strongly Disagree**	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
1. The assessor was genuinely interested in helping us.					
2. I have lots of new ideas about how to parent my child.					
3. My child felt comfortable with the assessor.					
4. My child's problems are partly caused by other struggles in our family.					
5. I felt the assessor respected me.					
6. My child never really warmed up to the assessor.					
7. I was informed about each step of the assessment.					
8. Many of my child's difficulties have to do with our family.					
9. I learned a tremendous amount about my child from this assessment.					
10. The assessment made me feel ashamed.					
11. I liked the assessor.					
12. The assessment revealed how family members play a role in my child's problems.					
13. I trusted the assessor.					
14. I felt blamed for my child's problems.					
15. I am better able to communicate with my child.					
16. I now see how our family's problems affect my child.					
17. My child and the assessor really connected well.					
18. The assessment made me feel like a bad parent.					
19. Now I know what to expect from my child.					
20. I felt judged by the assessor.					
21. I felt that my opinion was valued.					
22. My child did not like the assessor.					
23. The assessor really listened to me.					
24. I understand my child so much better now.					

References

Ackerman, S. J., Hilsenroth, M. J., Baity, M.R. & Blagys, M. D. (2000). Interaction of theapeutic process and alliance during psychological assessment. *Journal of Personality Assessment, 73,* 82-109.

Attkisson, C. C., & Greenfield, T. K. (2004). The UCSF client satisafction scales: I. The client satisfaction questionnaire. In M. E. Maruish (Ed.), *The Use of Psychological Testing for Treatment Planning and Oucomes Assessment* (Vol. 3, pp. 799-811). Lawrence Erlbaum Associates.

Attkisson, C. C., & Zwick, R. (1982). Pscyhometric properties and correlations with service utilization and pscyhotherapy outcome. *Evaluation and Program Planning, 5*, 233-237.

Bailey, D. B., & Simeonsson, R. J. (1988). *Family assessment in early intervention.* Columbus, OH: Merrill.

Bodin, D., Beetar, R., Yeates, K. O., Boyer, K., Colvin, A. N., & Mangeot, S. (2007). A survey of parent satisfaction with pediatric neuropsychological evaluation. *The Clinical Neuropsychologist*, 884-898.

Brannan, A. M., Sonnichsen, S. E., & Heflinger, C. A. (1996). Measuring satisfaction with children's mental health services: Validity and reliability of the satisfaction scales. *Education and Program Planning, 19* (2), 131-141.

Byalin, K. (1993). Assessing parental satisfaction with children's mental health services. *Evaluation and Program Planning, 16*, 69-72.

Byrne, B. M., & Stewart, S. M. (2006). The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling, 13* (2), 287-321.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16-29.

Crowne, D.P., & Marlow, D. (1961). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

DeVellis, R. F. (2003). *Scale development theory and applications* (Second ed.). Sage Publications, Inc.

Donabedian, A. (1988). The quality of care: How can it be assessed? *Journal of the American Medical Association, 260*, 1743-1748.

Essex, D. W., Fox, J. A., & Groom, J. M. (1981). The development, factor analysis, and revision of a client satisfaction form. *Community Mental Health Journal, 17* (3), 226-235.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of explorartory factor analysis in psychological research. *Psychological Methods, 4* (3), 272-299.

Fantuzzo, J., Perry, M. A., & Childs, S. (2006). Parent satisfaction with education experiences scale: A multivariate examination of parent satisfaction with early childhood education programs. *Early Childhood Research Quarterly, 21*, 142-152.

Finn, S. E. (1997, March). Collaborative child assessment as a family systems intervention. *Paper presented at the annual meeting of the Society for Personality Assessment.* San Diego, CA.

Finn, S. E. (2007). *In Our Client's Shoes: Theory and Techniques of Therapeutic Assessment.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Finn, S. E. (1996). *Manual for using the MMPI-2 as a therapeutic intervention.* Minneapolis: University of Minnosota Press.

Finn, S. E., & Kamphuis, J. H. (2006). Therapeutic assessment with the MMPI-2. In J. Butcher, *MMPI-2: A practitioner's guide* (pp. 165-191). Washington, D.C.: APA Books.

Finn, S.E., Schroeder, D.G., & Tonsager, M.E. (2004). *The Assessment Questionnaire-2 (AQ-2): A measure of clients' experiences with psychological assessment*. Unpublished manuscript.

Finn, S. E., & Tonsager, M. E. (2002). How thearpeutic assessment become humanistic. *Humanistic Psychologist, 30*, 10-22.

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment, 9*, 374-385.

Finn, S. E., & Tonsanger, M. E. (1992). The therapeutic effects of providing MMPI-2 feedback to college students awaiting pscyhotherapy. *Psychological Assessment, 4*, 278-287.

Finn, S. (2003). Therapeutic assessment of a man with "ADD". *Journal of Personality Assessment, 80*, 115-129.

Fischer, C. T. (1985/1994). *Individualizing psychological assessment.* Mahwah, NJ: Lawrence Erlbaum Associates.

French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling, 13* (3), 378-402.

Garland, A. F., Haine, R. A., & Boxmeyer, C. L. (2007). Determinates of youth and parent satisafction in usual care pscyhotherapy. *Evaluation and Program Planning, 30*, 45-54.

Gaston, L., & Sabourin, S. (1992). Client satisfaction and social desirability in psychiatry. *Evaluation and Program Planning, 15*, 227-231.

Gerkensmeyer, J. E., & Austin, J. K. (2005). Development and testing of a scale measuring parent satisfaction with staff interactions. *The Journal of Beahvioral Health Sevices & Research, 32* (1), 61-73.

Gerkensmeyer, J. E., Austin, J. K., & Miller, T. K. (2006). Model testing: Examining parent satisfaction. *Archives of Pscyhiatric Nursing, 20* (2), 65-75.

Graham, J.W. (2009) Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549-576.

Greenfield, T. K., & Attkisson, C. C. (1989). Steps toward a multifactorial satisfaction scale for primary care and mental health services. *Evaluation and Program Planning, 12*, 271-278.

Greenfield, T. K., & Attkisson, C. C. (2004). The UCSF Client Satisfaction Scales: II. The service satisafction score-30. In M. E. Maruish (Ed.), *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment* (Vol. 3, pp. 813-837). Lawrence Erlbaum Associates.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Pscyhological Bulletin, 103* (2), 265-275.

Hamilton, A., Fowler, J., Hersh, B., Austin, C. A., Finn, S. E., Tharinger, D. J., et al. (2009). The drama of the gifted child revaled through Therapeutic Assessment. *Journal of Personality Assessment* .

Handler, L. (2007). Therapeutic assessment with children and adolescents. In S. S. Handler, *Clinical asssessment of children and adolescents: A practioner's guide.* Mahwah, NJ: Lawrence Erlbaum Associates.

Harrington Godley, S., Fiedler, E. M., & Funk, R. R. (1998). Consumer satisfaction of parents and their children with child/adolescent mental health services. *Evaluation and Program Planning, 21*, 31-45.

Hu, L. & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Huebner, R.A., Jones, B. L., Miller, V. P., Custer, M., & Critchfield, B. (2006). Comprehensive family services and customer satisfaction outcomes. *Child Welfare, LXXXV*, (4), 691-714.

Kapp, S. A., & Vela, R. H. (2004). The parent satisfaction with foster care services scale. *Child Welfare, LXXXIII* (3), 263-287.

Kaufman, C., & Phillips, D. (2000). *Survey of state consumer surveys.* Rockville, MD: SAMHSA.

Keith, T. Z. (2006). *Multiple regression and beyond.* Pearson Education, Inc.

King, G., Cathers, T., King, S., & Rosenbaum, P. (2001). Major elements of parents' satisfaction and dissatisfaction with pediatric rehabilitation services. *Children's Health Care, 30* (2), 111-134.

Kline, P. (1994). *An easy guide to factor analysis.* Routledge.

Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2*, 197-207.

LaSala, M. C. (1997). Client satisfaction: Consideration of correlated and response bias. *Familiies in Society*, 54-64.

Lebow, J. (1982). Consumer satisfaction with mental health treatment. *Pscyhological Bulletin, 91* (2), 244-259.

Lewis, J. R. (1994). Patient views on quality care in general practice: Literature review. *Social Science Medicine, 39* (5), 655-670.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.

Martin, J. S., Peter, C. G., & Kapp, S. A. (2003). Consumer satisfaction with children's mental health services. *Child and Adolescent Social Work Journal, 20*, 211-226.

McNaughton, D. (1994). Measuring parent satisfaction with early childhood intervention programs: Current practice, problems, and future perspectives. *Topics in Early Childhood Special Education, 14* (1), 26-48.

Measelle, J. R., Weinstein, R. S., & Martinez, M. (1998). Parent satisfaction with case managed systems of care for children and youth with severe emotional disturbance. *Journal of Chld and Family Studies, 7* (4), 451-467.

Pascoe, G. C. (1984). Patient satisfaction in primary health care: A literature review and analysis. *Evaluation and Program Planning, 6*, 185-210.

Plante, T. G., Couchman, C. E., & Hoffman, C. A. (1998). Measuring treatment outcome and client satisfaction among children and families: A case report. *Professional Psychology: Research and Practice, 29* (1), 32-55.

Rey, J. M., O'Brien, M., & Walter, G. (2002). Is the satisfied customer one who also 'does well'? The relationship between outcome and parent satisfaction in child and adolescent mental health service. *Australian Psychiatry, 10* (3), 246-249.

Reynolds, M. R., & Keith, T. Z. (2009). Measurement and statistical issues in child assessment research. In C. R. Reynolds (Ed.), *Oxford handbook of pscyhological assessment of children and adolescents.* New York: Oxford University.

Reynolds, M. R., Keith, T. Z., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent general and broad cognitive abilities for children and youth: Evidence from higher order MG-MACS and MIMIC models. *Intelligence, 36*, 236-260.

Riley, S. E., Stromberg, A. J., & Clark, J. (2005). Assessing prental satisfaction with children's mental health services with the Youth Services Survey for Families. *Journal of Child and Family Studies, 14* (1), 87-99.

Robertson, A.S. (2006). Including parents, foster parents and parenting caregivers in the assessments and interventions of young children placed in the foster care system. *Children and Youth Services Review, 28*, 180-192.

Shapiro, J. P., Welker, C. J., & Jacobson, B. J. (1997). The Youth Client Satisfaction Questionnaire: Development, construct validation, and factor structure. *Journal of Child Clinical Psychology, 26* (1), 87-98.

Sheppard, M. (1993). Client satisfaction, extended intervention and interpersonal skills in community mental health. *Journal of Advanced Nursing, 18*, 246-259.

Stallard, P. (2001). Reducting parental dissatisfaction with a child and adolescent psychology service: A process of quality improvement. *Journal of Mental Health, 10* (1), 63-73.

Summers, J. A., Hoffman, L., Marquis, J., Turnbull, A., Poston, D., & Nelson, L. L. (2005). Measuring the quality of family-professional partnerships in special education services. *Exceptional Children, 72* (1), 65-81.

Tellegen, A. (1981). Practicing the two disciplines for relaxation and enlightenment: Comment on "Role of the feddback signfoar in electromyograph biodfeedback: The relevance of attention" by Qualls and Sheehan. *Journal of Experimental Psychology, 110* (2), 217-226.

Tharigner, D. J., Finn, S. E., Wilkinson, A., & Schaber, P. (2007). Therapeutic Assessment with children as a family intervention: Clinical protocol and a research case study. *Psychology in Schools, 44*, 293-309.

Tharigner, D. J., Finn, S. E., Wilkinson, A., DeHay, T., Parton, V. T., Bailey, K. E., et al. (2008). Providing Psychological assessment feedback with children through indvidualized fables. *Professional Psychology: Research and Practice* .

Tharinger, D. J., Finn, S. E., Arora, P., Ihorn, S., Judd-Glossy, L., & Wan, J. (in preparation). Parental observation of the children's testing sessions in Therapeutic Assessment: Process and impact.

Tharinger, D. J., Finn, S. E., Austin, C. A., Gentry, L. B., Bailey, K. E., Parton, V. T., et al. (2008). Family Sessions in psychological assessment with children: Goals, techniques, and clinical utility. *Journal of Personality Assessment* .

Tharinger, D. J., Finn, S. E., Gentry, L. B., Hamilton, A., Fowler, J., Matson, M., et al. (2009). Therapeutic Assessment with children: A pilot study of treatment acceptability and outcome. *Journal of Personality Assessment, 90*, 547-558.

Tharinger, D. J., Finn, S. E., Hersh, B., Wilkinson, A., Christopher, G., & Tran, A. (2008). Assessment feedback with parents and pre-adolescent children: A collaborative approach. *Professional Practice: Research and Practice* .

Tharinger, D. J., Krumholz, L., Austin, C. A., & Matson, M. (In press). The development and model of therapeutic assessment with children: Application to school-based assessment. In M. A. Bray, & T. J. Kehle (Eds.), *Oxford Press Handbook of School Psychology.* Oxford University Press.

Williams, B., Coyle, J., & Healy, D. (1998). The meaning of patient satisfaction: An explanation of high reported levels. *Social Science Medicine, 47* (9), 1351-1359.

Williams, S. J., & Calnan, M. (1991). Key detereminants of consumer satisfaction with general practice. *Family Practice, 8*, 237-242.

Young, S. C., Nicholson, J., & Davis, M. (1995). An overview of issues in research on consumer satisfaction with child and adolescent mental health services. *Journal of Child and Family Studies, 4* (2), 219-238.